

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2004, 12:59:25 ; Search time 21 seconds
(without alignments)
556.086 Million cell updates/sec

Title: US-09-880-503-5

Perfect score: 1508

Sequence: 1 KPSSPPEELKFQCGQKTLRP.....VSHFLPWIRSHTEKENGAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PGTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1508	100.0	365	1	US-08-093-741-83
2	1508	100.0	365	1	US-08-720-012-83
3	1508	100.0	393	2	US-08-560-098A-44
4	1508	100.0	393	3	US-08-967-024C-24
5	1508	100.0	393	3	US-08-967-024C-25
6	1508	100.0	411	1	US-08-087-163-1
7	1508	100.0	411	1	US-08-286-748B-18
8	1508	100.0	411	1	US-08-153-799-18
9	1508	100.0	411	3	US-09-181-816-1
10	1508	100.0	430	6	5219569-2
11	1508	100.0	431	4	US-09-101-272G-1
12	1508	100.0	431	6	5188829-1
13	1508	100.0	432	2	US-08-560-098A-47
14	1508	99.8	430	1	US-07-942-157A-3
15	1496	99.2	306	2	US-08-560-098A-45
16	1496	99.2	331	2	US-08-560-098A-46
17	1495	99.1	411	2	US-08-560-098A-48
18	1382	91.6	253	3	US-08-944-483-73
19	1374	91.1	254	2	US-08-560-098A-49
20	607	40.3	355	2	US-08-811-949-59
21	590	39.1	355	2	US-08-811-949-47
22	590	39.1	437	2	US-08-811-949-51
23	590	39.1	527	2	US-08-811-949-39
24	587	38.9	437	2	US-08-811-949-57
25	583	38.7	355	1	US-08-137-116-1
26	583	38.7	355	1	US-08-217-618-1
27	583	38.7	355	1	US-08-427-640-2

28	583	38.7	355	1	US-08-427-640-6	Sequence 6, Appli
29	583	38.7	355	1	US-08-217-617A-1	Sequence 1, Appli
30	583	38.7	355	1	US-08-217-616-1	Sequence 1, Appli
31	583	38.7	355	2	US-08-811-949-45	Sequence 45, Appli
32	583	38.7	355	3	US-08-794-528-1	Sequence 1, Appli
33	583	38.7	355	6	5223256-1	Patent No. 5223256
34	583	38.7	356	1	US-08-427-640-4	Sequence 4, Appli
35	583	38.7	356	1	US-08-427-640-8	Sequence 8, Appli
36	583	38.7	378	4	US-09-553-498-10	Sequence 10, Appli
37	583	38.7	378	4	US-09-618-869-10	Sequence 10, Appli
38	583	38.7	383	2	US-08-558-269-6	Sequence 6, Appli
39	583	38.7	383	3	US-09-410-882-6	Sequence 6, Appli
40	583	38.7	472	2	US-08-811-949-63	Sequence 63, Appli
41	583	38.7	527	1	US-07-609-510B-16	Sequence 16, Appli
42	583	38.7	527	5	PCT-US91-01025A-2	Sequence 2, Appli
43	583	38.7	527	6	5155259-8	Patent No. 5155259
44	583	38.7	562	2	US-08-811-949-43	Sequence 43, Appli
45	583	38.7	562	2	US-08-560-098A-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDEL, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,741
; FILING DATE: 20-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-093-741-83

Query Match 100.0%; Score 1508; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.8e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCKQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGSGSVTVYCGSLIS 60
 Db 90 KPSSPPEELKFCQCKQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGSGSVTVYCGSLIS 149
 Qy 61 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
 Db 150 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 209
 Qy 121 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 180
 Db 210 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 269
 Qy 181 VVKLIHSHRCQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 270 VVKLIHSHRCQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 2
 US-08-720-012-83
 ; Sequence 83, Application US/08720012
 ; Patent No. 5747291
 ; GENERAL INFORMATION:
 ; APPLICANT: STEFFENS, GERD J.
 ; APPLICANT: WENDEL, STEPHAN
 ; APPLICANT: SCHNEIDER, JOHANNES
 ; APPLICANT: HEINZEL-WIELAND, REGINA
 ; APPLICANT: SAUNDERS, DEREK J.
 ; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
 ; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
 ; TITLE OF INVENTION: INHIBITING EFFECT
 ; NUMBER OF SEQUENCES: 83
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N. W. Suite 700
 ; CITY: Washington, D.C.
 ; COUNTRY: U.S.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/720,012
 ; FILING DATE: 27-SEP-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/093,741
 ; FILING DATE: 20-JUL-1993
 ; APPLICATION NUMBER: DE P43 23 754.1
 ; FILING DATE: 15-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/41345
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 83:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 365 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-720-012-83

Query Match 100.0%; Score 1508; DB 1; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1.8e-148;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCKQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGSGSVTVYCGSLIS 60
 Db 90 KPSSPPEELKFCQCKQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGSGSVTVYCGSLIS 149
 Qy 61 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
 Db 150 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 209
 Qy 121 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 180
 Db 210 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 269
 Qy 181 VVKLIHSHRCQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 270 VVKLIHSHRCQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 3
 US-08-560-098A-44
 ; Sequence 44, Application US/08560098A
 ; Patent No. 5976841
 ; GENERAL INFORMATION:
 ; APPLICANT: WENDEL, STEPHAN
 ; APPLICANT: HEINZEL-WIELAND, REGINA
 ; APPLICANT: STEFFENS, GERD JOSEF
 ; TITLE OF INVENTION: Proteins Having Fibrinolytic and
 ; TITLE OF INVENTION: Coagulation-Inhibiting Properties
 ; NUMBER OF SEQUENCES: 60
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,098A
 ; FILING DATE: 17-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 40 892.7
 ; FILING DATE: 17-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42448
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 44:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 393 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-560-098A-44

Query Match 100.0%; Score 1508; DB 2; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2e-148;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCKQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGSGSVTVYCGSLIS 60
 Db 91 KPSSPPEELKFCQCKQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGSGSVTVYCGSLIS 150

Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 210
 Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
 Db 211 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 270
 Qy 181 VKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 271 VKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 330
 Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
 Db 331 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 366

RESULT 4
 US-08-967-024C-24
 ; Sequence 24, Application US/08967024C
 ; Patent No. 6133011
 ; GENERAL INFORMATION:
 ; APPLICANT: WNENDT, Stephan
 ; APPLICANT: STEFFENS, Gerd Josef
 ; APPLICANT: JANOSCH, Elke
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/967,024C
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 42 665.8
 ; FILING DATE: 30-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42444
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 393 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-967-024C-24

Query Match 100.0%; Score 1508; DB 3; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2e-148;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 60
 Db 91 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 150
 Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 210
 Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180

Db 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 210
 Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
 Db 211 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 270
 Qy 181 VKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 271 VKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 330
 Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
 Db 331 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 366

RESULT 5
 US-08-967-024C-25
 ; Sequence 25, Application US/08967024C
 ; Patent No. 6133011
 ; GENERAL INFORMATION:
 ; APPLICANT: WNENDT, Stephan
 ; APPLICANT: STEFFENS, Gerd Josef
 ; APPLICANT: JANOSCH, Elke
 ; APPLICANT: HEINZEL-WIELAND, Regina
 ; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 ; STREET: 1200 G Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: DC
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/967,024C
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: P 44 42 665.8
 ; FILING DATE: 30-NOV-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: EVANS, Joseph D.
 ; REGISTRATION NUMBER: 26,269
 ; REFERENCE/DOCKET NUMBER: 148/42444
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-8800
 ; TELEFAX: (202) 628-8844
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 393 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-967-024C-25

Query Match 100.0%; Score 1508; DB 3; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2e-148;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 60
 Db 91 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 150
 Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 210
 Qy 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180

Db 211 IALLKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGKENSSTLYLPEQLKMT 270
 Qy 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 271 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 330
 Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 331 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 366
 RESULT 6
 US-08-087-163-1
 ; Sequence 1, Application US/08087163
 ; Patent No. 5472692
 ; GENERAL INFORMATION:
 ; APPLICANT: Gurewich, Victor
 ; TITLE OF INVENTION: PRO-UKINASE MUTANTS
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/087,163
 ; FILING DATE: 07/02/93
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fasse, J. Peter
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 04353/003001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411
 ; TYPE: amino acid
 ; STRANDEDNESS: N/A
 ; TOPOLOGY: N/A
 ; US-08-087-163-1
 Query Match 100.0%; Score 1508; DB 1; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-148;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYCGSLIS 60
 Db 136 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYCGSLIS 195
 Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 Qy 121 IALLKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGKENSSTLYLPEQLKMT 180
 Db 256 IALLKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGKENSSTLYLPEQLKMT 315
 Qy 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 316 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 376 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
 RESULT 7
 US-08-286-748B-18
 ; Sequence 18, Application US/08286748B
 ; Patent No. 5759542
 ; GENERAL INFORMATION:
 ; APPLICANT: Victor Gurewich
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
 ; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 502 or 55SX
 ; OPERATING SYSTEM: MS-DOS (Version 5.0)
 ; SOFTWARE: WordPerfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/286,748B
 ; FILING DATE: August 5, 1994
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: J. Peter Fasse
 ; REGISTRATION NUMBER: 32,983
 ; REFERENCE/DOCKET NUMBER: 04547/013001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-286-748B-18
 Query Match 100.0%; Score 1508; DB 1; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-148;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYCGSLIS 60
 Db 136 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYCGSLIS 195
 Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 Qy 121 IALLKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGKENSSTLYLPEQLKMT 180
 Db 256 IALLKRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGKENSSTLYLPEQLKMT 315
 Qy 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 316 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 8
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; APPLICATION DATA: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 1508; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.2e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGQKTLPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 60
DB 136 KPSSPPEELKFCQGQKTLPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 195

QY 61 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 196 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 180
DB 256 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 315

QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPWKTDSCGDSGGPLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECCQPHYGVSEVTTKMLCAADPWKTDSCGDSGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 9
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 100.0%; Score 1508; DB 3; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.2e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGQKTLPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 60
DB 136 KPSSPPEELKFCQGQKTLPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 195

QY 61 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 196 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 180
DB 256 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFKENSTDYLYPEQLKMT 315

QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPWKTDSCGDSGGPLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECCQPHYGVSEVTTKMLCAADPWKTDSCGDSGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 10
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2
; LENGTH: 430
5219569-2

Query Match 100.0%; Score 1508; DB 6; Length 430;

```

Best Local Similarity 100.0%; Pred. No. 2.3e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPELKFQCGQKTLRPRFKIIGGFTTIEHQWFAAIYRHRGSGVTVYVCGSLIS 60
Db 155 KPSSPPELKFQCGQKTLRPRFKIIGGFTTIEHQWFAAIYRHRGSGVTVYVCGSLIS 214
QY 61 PCWVISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 215 PCWVISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 274
QY 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
Db 275 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
QY 181 VKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 335 VKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 394
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
Db 395 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 430

RESULT 11
US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; FILE REFERENCE: Q50979
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

Query Match
Best Local Similarity 100.0%; Score 1508; DB 4; Length 431;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPELKFQCGQKTLRPRFKIIGGFTTIEHQWFAAIYRHRGSGVTVYVCGSLIS 60
Db 156 KPSSPPELKFQCGQKTLRPRFKIIGGFTTIEHQWFAAIYRHRGSGVTVYVCGSLIS 60
QY 61 PCWVISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
QY 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
QY 181 VKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 336 VKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 431

RESULT 13
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269

```

```

Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 431

RESULT 12
518829-1
; Patent No. 518829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAHO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO:1
; LENGTH: 431
518829-1

Query Match
Best Local Similarity 100.0%; Score 1508; DB 6; Length 431;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPELKFQCGQKTLRPRFKIIGGFTTIEHQWFAAIYRHRGSGVTVYVCGSLIS 60
Db 156 KPSSPPELKFQCGQKTLRPRFKIIGGFTTIEHQWFAAIYRHRGSGVTVYVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
QY 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
QY 181 VKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 336 VKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 431

RESULT 13
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269

```

```

; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-47

Query Match 100.0%; Score 1508; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.3e-148;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVVCGSLIS 60
DB 157 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVVCGSLIS 216
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 217 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 276
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
DB 277 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMT 336
QY 181 VVKLISHRECCQPHYVSEVTTKMLCAADPOWKTDCQDGGSGPLVCSLOGRMTLTGIVS 240
DB 337 VVKLISHRECCQPHYVSEVTTKMLCAADPOWKTDCQDGGSGPLVCSLOGRMTLTGIVS 396
QY 241 WGRGCALKDKPGVYTRVSHFLPWRSHRSHKEENGLAL 276
DB 397 WGRGCALKDKPGVYTRVSHFLPWRSHRSHKEENGLAL 432

RESULT 14
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label=peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label=modified
; OTHER INFORMATION: /note="six amino acids deleted in mutant"
; US-07-942-157A-3

Query Match 99.8%; Score 1505; DB 1; Length 430;
Best Local Similarity 99.6%; Pred. No. 4.7e-146;
Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVVCGSLIS 60
DB 155 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVVCGSLIS 214
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 215 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 274
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
DB 275 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMT 334
QY 181 VVKLISHRECCQPHYVSEVTTKMLCAADPOWKTDCQDGGSGPLVCSLOGRMTLTGIVS 240
DB 335 VVKLISHRECCQPHYVSEVTTKMLCAADPOWKTDCQDGGSGPLVCSLOGRMTLTGIVS 394
QY 241 WGRGCALKDKPGVYTRVSHFLPWRSHRSHKEENGLAL 276
DB 395 WGRGCALKDKPGVYTRVSHFLPWRSHRSHKEENGLAL 430

RESULT 15
US-08-560-098A-45
; Sequence 45, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:

```

NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 306 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-098A-45

Query Match 99.2%; Score 1496; DB 2; Length 306;
 Best Local Similarity 100.0%; Pred. No. 2.5e-147;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 SSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLISPC 62
 DB 33 SSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLISPC 92
 QY 63 WVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIA 122
 DB 93 WVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIA 152
 QY 123 LKIRSKGRCQAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPQLKMTVV 182
 DB 153 LKIRSKGRCQAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPQLKMTVV 212
 QY 183 KLIHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSWG 242
 DB 213 KLIHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSWG 272
 QY 243 RGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 273 RGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 306

RESULT 16
 US-08-560-098A-46
 Sequence 46, Application US/08560098A
 Patent No. 5976841
 GENERAL INFORMATION:
 APPLICANT: WENNDT, Stephan
 APPLICANT: HEINZEL-WIELAND, Regina
 APPLICANT: STEFFENS, Gerd Josef
 TITLE OF INVENTION: Proteins having Fibrinolytic and
 TITLE OF INVENTION: Coagulation-inhibiting Properties
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 STREET: 1200 G Street, N.W., Suite 700
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560,098A
 FILING DATE: 17-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P 44 40 892.7
 FILING DATE: 17-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 46:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 331 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-098A-46

Query Match 99.2%; Score 1496; DB 2; Length 331;
 Best Local Similarity 100.0%; Pred. No. 2.8e-147;
 Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 SSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLISPC 62
 DB 33 SSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLISPC 92
 QY 63 WVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIA 122
 DB 93 WVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIA 152
 QY 123 LKIRSKGRCQAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPQLKMTVV 182
 DB 153 LKIRSKGRCQAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPQLKMTVV 212
 QY 183 KLIHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSWG 242
 DB 213 KLIHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSWG 272
 QY 243 RGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 273 RGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 306

RESULT 17
 US-08-560-098A-48
 Sequence 48, Application US/08560098A
 Patent No. 5976841
 GENERAL INFORMATION:
 APPLICANT: WENNDT, Stephan
 APPLICANT: HEINZEL-WIELAND, Regina
 APPLICANT: STEFFENS, Gerd Josef
 TITLE OF INVENTION: Proteins having Fibrinolytic and
 TITLE OF INVENTION: Coagulation-inhibiting Properties
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
 STREET: 1200 G Street, N.W., Suite 700
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/560,098A
 FILING DATE: 17-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P 44 40 892.7
 FILING DATE: 17-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:
LENGTH: 411 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-48

Query Match 99.1%; Score 1495; DB 2; Length 411;

Best Local Similarity 99.3%; Pred. No. 4.9e-147; Indels 0; Gaps 0;
Matches 274; Conservative 0; Mismatches 2;

QY 1 KPSSPEELKFCQGKTLRPRKIIIGSEFTTIENQPFPAAIYRRHRGGSVTVCGSLIS 60
DB 136 KPSSPEELKFCQGKTLRPRKIIIGSEFTTIENQPFPAAIYRRHRGGSVTVCGSLIS 195
QY 61 PCWVISATCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENILHKDYSADTLAHND 120
DB 196 PCWVISATCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENILHKDYSADTLAHND 255
QY 121 IALLKRSKEGRCQAQPSRTIQICLPSMYNDPOFGTSCITGFGKENSTDYLYPEQLKMT 180
DB 256 IALLKRSKEGRCQAQPSRTIQICLPSMYNDPOFGTSCITGFGKENSTDYLYPEQLKMT 315
QY 181 VKLISHRECOQPHYGSEVTTIMLCAADPOWKTSCQDGGPLVCSLQCRMTLTGIVS 240
DB 316 VKLISHRECOQPHYGSEVTTIMLCAADPOWKTSCQDGGPLVCSLQCRMTLTGIVS 375
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 376 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLVL 411

RESULT 18

US-08-944-483-73

Sequence 73, Application US/08944483

Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLASS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.

APPLICANT: STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

TITLE OF INVENTION: OF THE PROSTATE

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 73:

SEQUENCE CHARACTERISTICS:

LENGTH: 253 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6232456e

US-08-944-483-73

Query Match 91.6%; Score 1382; DB 3; Length 253;

Best Local Similarity 100.0%; Pred. No. 1.4e-135; Indels 0; Gaps 0;

Matches 253; Conservative 0; Mismatches 0;

QY 24 IIGSEFTTIENQPFPAAIYRRHRGGSVTVCGSLISPCWVISATHCFIDYPKKEDYIVY 83

DB 1 IIGSEFTTIENQPFPAAIYRRHRGGSVTVCGSLISPCWVISATHCFIDYPKKEDYIVY 60

QY 84 LGRRLNSNTQGMKEFEVENILHKDYSADTLAHNDIALLKIRSEGRCAQPSRTTQTI 143

DB 61 LGRRLNSNTQGMKEFEVENILHKDYSADTLAHNDIALLKIRSEGRCAQPSRTTQTI 120

QY 144 CLPSMYNDPOFGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTK 203

DB 121 CLPSMYNDPOFGTSCITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGSEVTTK 180

QY 204 MLCAADPOWKTSCQDGGPLVCSLQCRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 263

DB 181 MLCAADPOWKTSCQDGGPLVCSLQCRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 240

QY 264 IRSHTKEENGLAL 276

DB 241 IRSHTKEENGLAL 253

RESULT 19

US-08-560-098A-49

Sequence 49, Application US/08560098A

Patent No. 5978841

GENERAL INFORMATION:

APPLICANT: WNEBET, Stephan

APPLICANT: HEINZEL-WIELAND, Regina

APPLICANT: STEPFENS, Gerd Josef

TITLE OF INVENTION: Proteins having Fibrinolytic and

TITLE OF INVENTION: Coagulation-inhibiting Properties

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N.W., Suite 700

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/560,098A

FILING DATE: 17-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: P 44 40 892.7

FILING DATE: 17-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42448

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-49

Query Match 91.1%; Score 1374; DB 2; Length 254;
Best Local Similarity 99.2%; Pred. No. 9.2e-135;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 23 KIIGFTTIENQPFPAIYRHRGSGVTYVCGSLISPCWVISATHCFIDYPPKEDYIV 82
DB 1 KIIGFTTIENQPFPAIYRHRGSGVTYVCGSLISPCWVISATHCFIDYPPKEDYIV 60
QY 83 YLGRRLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQT 142
DB 61 YLGRRLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQT 120
QY 143 ICLPSMYNDPQFGTSCETITGFGKENSVDLYPEQLKMTVVKLIHSHRECQPHYGVSEVTT 202
DB 121 ICLPSMYNDPQFGTSCETITGFGKENSVDLYPEQLKMTVVKLIHSHRECQPHYGVSEVTT 180
QY 203 KVLCAADPOWKDSCQDGGPLVCSLOGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLP 262
DB 181 KVLCAADPOWKDSCQDGGPLVCSLOGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLP 240
QY 263 WIRSHKTEENGLAL 276
DB 241 WIRSHKTEENGLVL 254

RESULT 20

US-08-811-949-59
; Sequence 59, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-811-949-59

Query Match 40.3%; Score 607; DB 2; Length 355;
Best Local Similarity 46.8%; Pred. No. 7.6e-55;
Matches 124; Conservative 36; Mismatches 95; Indels 10; Gaps 5;

QY 13 CG-QKTLPFRFKIIGBEFTTIENQPFPAIYRHR-RGGSVTYVCGSLISPCWVISATHC 70
DB 92 CGLRQTLPFRFKIIGBEFTTIENQPFPAIYRHRSGRFLCGGILISSCWLSAHC 151
QY 71 FIDYPKKEDYIVYVGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSK 130
DB 152 FQERFPFHLLTVILGRTYRVVPGEEQKFEVEYIVHKFDDDT--YDNDIALQLKSDS 209
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSVDLYPEQLKMTVVKLIHSHREC 190
DB 210 SRCAQESSVVRTVCLPPADQLDPDTECELSGFKHEALSPFYSERLKEAHRVLYPSSRC 269
QY 191 QQPHYGVSEVTTKMLCAAD-----PQWKT-DSCQDGGSGPLVCSLOGRMTLTGIVSWGRG 244
DB 270 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDA CQDGGSGPLVCLNDGRMTLVGLISWGLG 329
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 330 CGQKDFGVYTKVTNYLDWIRDNMR 354

RESULT 21

US-08-811-949-47
; Sequence 47, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-47

Query Match 39.1%; Score 590; DB 2; Length 355;
Best Local Similarity 45.3%; Pred. No. 4.4e-53;
Matches 120; Conservative 39; Mismatches 96; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKIIGGEFTTIENOPFAAIYRRH-RGGSVTVVCGGSLISPCWISATHC 70
DB 92 CGLRQYSQPFRIIGLGFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151
QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 152 FQERFPHHLTVILGRYRVVPGEEQKFEVEKYVHKFDDDT--YNDIALQLKSDS 209
QY 131 GRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 210 SRCAQESSVVRTVCLPADLQPDWTECELSGYGKHEALSPFYSERLKEAHLVLPSSRC 269
QY 191 QPPHYVGSEVTTKMLCAAD-----PWKT-DSOQDGGPLVCSLOGRMTLTGIVSWG 244
DB 270 TSQHLLNRTVTDNMLCAGDTRSGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 330 CGQKDVPGVYTKVNYLDWIRDNR 354

RESULT 22

US-08-811-949-51
Sequence 51, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-811-949-51

Query Match 39.1%; Score 590; DB 2; Length 437;
Best Local Similarity 45.3%; Pred. No. 6e-53;
Matches 120; Conservative 39; Mismatches 96; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKIIGGEFTTIENOPFAAIYRRH-RGGSVTVVCGGSLISPCWISATHC 70
DB 174 CGLRQYSQPFRIIGLGFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 233
QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 234 FQERFPHHLTVILGRYRVVPGEEQKFEVEKYVHKFDDDT--YNDIALQLKSDS 291
QY 131 GRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 292 SRCAQESSVVRTVCLPADLQPDWTECELSGYGKHEALSPFYSERLKEAHLVLPSSRC 351
QY 191 QPPHYVGSEVTTKMLCAAD-----PWKT-DSOQDGGPLVCSLOGRMTLTGIVSWG 244
DB 352 TSQHLLNRTVTDNMLCAGDTRSGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 411
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 412 CGQKDVPGVYTKVNYLDWIRDNR 436

RESULT 23

US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39
Query Match 39.1%; Score 590; DB 2; Length 527;

Best Local Similarity 45.3%; Pred. No. 7.8e-53;
Matches 120; Conservative 39; Mismatches 96; Indels 10; Gaps 5;
13 CG-QKTLRPRFKIIGBFTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
174 CGLRQYSQFQDIIGGLFADIASHPQAAIPAKRRSPGERFLCGGILISSCWILSAHC 233
71 FIDVPKEDYIVVLCGRSLNSNTQGMKFEVENILHKDYSADTLAHHNDIALKIRSK 130
234 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKIVHKEFDDT--YNDNDIALQLKSDS 291
131 GRCAQPSRTIOTICLPSMYNDPQGTSCETITGPKENSTDYLYPEOLKMTVVKLISHREC 190
292 SRCQAGSSVVRTVCLPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRRC 351
191 QOPHYGSEVTTKMLCAAD-----PWKT-DSQCQSDSGPLVCSLQGRMTLTGIVSWG 244
352 TSQHLLNRTVTDNMLCAGDTRSGGPOANLHDACQSDSGPLVCLNDGRMTLVGIISWGLG 411
245 CALKDKPGVTVTRVSHFLPWIRSHTK 269
502 CGQKDVPGVTVKVTNYLDWIRDNR 526

SULT 24
-08-811-949-57
Sequence 57, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NINJA, MINO
APPLICANT: SAITO, YOSHIKAWA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: P.C.
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 437 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
-08-811-949-57

ery March 38.9%; Score 587; DB 2; Length 437;
Local Similarity 45.3%; Pred. No. 1.2e-52;
Matches 120; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGBFTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
DB 174 CGLRQYSQFQDIIGGLFADIASHPQAAIPAKRRSPGERFLCGGILISSCWILSAHC 233
QY 71 FIDVPKEDYIVVLCGRSLNSNTQGMKFEVENILHKDYSADTLAHHNDIALKIRSK 130
DB 234 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKIVHKEFDDT--YNDNDIALQLKSDS 291
QY 131 GRCAQPSRTIOTICLPSMYNDPQGTSCETITGPKENSTDYLYPEOLKMTVVKLISHREC 190
DB 292 SRCQAGSSVVRTVCLPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRRC 351
QY 191 QOPHYGSEVTTKMLCAAD-----PWKT-DSQCQSDSGPLVCSLQGRMTLTGIVSWG 244
DB 352 TSQHLLNRTVTDNMLCAGDTRSGGPOANLHDACQSDSGPLVCLNDGRMTLVGIISWGLG 411
QY 245 CALKDKPGVTVTRVSHFLPWIRSHTK 269
DB 412 CGQKDVPGVTVKVTNYLDWIRDNR 436
RESULT 25
US-08-137-116-1
; Sequence 1, Application US/08137116
; Patent No. 5500411
; GENERAL INFORMATION:
; APPLICANT: Martin, Ulrich
; APPLICANT: Koenig, Reinhard
; TITLE OF INVENTION: METHOD FOR TREATING THROMBOEMBOLIC
; TITLE OF INVENTION: CONDITIONS BY INHIBITING REOCCLUSION
; TITLE OF INVENTION: VIA THE USE OF MULTIPLE BOLUS
; TITLE OF INVENTION: ADMINISTRATION OF THROMBOLYTICALLY
; TITLE OF INVENTION: ACTIVE PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,116
; FILING DATE: June 30, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP92/00851
; FILING DATE: 15 April 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 41 12 398
; FILING DATE: 16 April 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 41 23 845
; FILING DATE: 18 July 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5500411man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: BOER 1026
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acids
; TOPOLOGY: linear
; US-08-137-116-1


```
Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 2.4e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFALYYRRH-RGGSVTYVCGSLISPCWVISAHC 70
DB 92 CGLRQYSQPQFRIKGLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151

QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLILHKDYSADTLAHNDIALKIRSK 130
DB 152 FQERFPFHLLTVILGRTRYRVVPGEEQKFEVKYIVHKEFDDDT--YDNDIALQLKSDS 209

QY 131 GRCAQPSRTIOTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 210 SRCQAESVVRVTVCLPADLQLPDWTCELSGKGHEALSPFYSERLKEAHRVLYPSSRC 269

QY 191 QPHYYGSEVTTKMLCAAD-----PQWKT-DSQCQDSGGPLVCSLQGRMTLTGIVSWG 244
DB 270 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329

QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 330 CGQKDVPGVYTKVNYLDWIRDNR 354

RESULT 26
US-08-217-618-1
; Sequence 1, Application US/08217618
; Patent No. 5510330
; GENERAL INFORMATION:
; APPLICANT: Martin, Ulrich
; TITLE OF INVENTION: COMBINATIONS OF THROMBOCYTICALLY ACTIVE
; TITLE OF INVENTION: PROTEINS AND NON-HEPARIN ANTICOAGULANTS, AND USES THEREOF
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/217,618
; FILING DATE: 25-MARCH-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5510330man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: BOER 1032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acids
; TOPOLOGY: linear
US-08-217-618-1

Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 2.4e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFALYYRRH-RGGSVTYVCGSLISPCWVISAHC 70
DB 92 CGLRQYSQPQFRIKGLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151

QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLILHKDYSADTLAHNDIALKIRSK 130
DB 152 FQERFPFHLLTVILGRTRYRVVPGEEQKFEVKYIVHKEFDDDT--YDNDIALQLKSDS 209

QY 131 GRCAQPSRTIOTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 210 SRCQAESVVRVTVCLPADLQLPDWTCELSGKGHEALSPFYSERLKEAHRVLYPSSRC 269

QY 191 QPHYYGSEVTTKMLCAAD-----PQWKT-DSQCQDSGGPLVCSLQGRMTLTGIVSWG 244
DB 270 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329

QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 330 CGQKDVPGVYTKVNYLDWIRDNR 354

RESULT 27
US-08-427-640-2
; Sequence 2, Application US/08427640
; Patent No. 5658788
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/427,640
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/689,410
; FILING DATE: 22 APRIL 1991
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-427-640-2

Query Match 38.7%; Score 583; DB 1; Length 355;
Best Local Similarity 44.9%; Pred. No. 2.4e-52;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFALYYRRH-RGGSVTYVCGSLISPCWVISAHC 70
DB 92 CGLRQYSQPQFRIKGLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 151

QY 71 FIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLILHKDYSADTLAHNDIALKIRSK 130
DB 152 FQERFPFHLLTVILGRTRYRVVPGEEQKFEVKYIVHKEFDDDT--YDNDIALQLKSDS 209

QY 131 GRCAQPSRTIOTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 210 SRCQAESVVRVTVCLPADLQLPDWTCELSGKGHEALSPFYSERLKEAHRVLYPSSRC 269

QY 191 QPHYYGSEVTTKMLCAAD-----PQWKT-DSQCQDSGGPLVCSLQGRMTLTGIVSWG 244
DB 270 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 329

QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
```

0 13:13:11 2004

Db

330 CQKXDPGVYTKVNTYLDNR 354
RESULT 28
US-08-427-640-6
Sequence 6, Application US/08427640
GENERAL INFORMATION:
APPLICANT: BIRG et al.
TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disorder
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS: 28
CITY: Lilly Lilly and Company
STATE: Indianapolis
COUNTRY: IN
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
OPERATING SYSTEM: Macintosh
CURRENT APPLICATION NUMBER: US/08/427,640
FILING DATE: 07/689,410
SEQUENCE CHARACTERISTICS: 6:
LENGTH: 355 amino acids
TOPOLOGY: linear
US-08-427-640-6

Query Match
Best Local Similarity 38.7%; Score 583; DB 1; Length 355;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
13 CG-KTLRPRFKLIGGEFTTINQWFAAIYRHH-RGGSVTVYCGSLSPCWVISAHC 70
92 CGLRQYQSPQRIKGGFLADIASHPWQAIFAKHRSRSPGERFLCGGILISSCHWLSAHC 151
71 FIDYPKEDYVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
152 FOERFPPLHLLTVILGRSRTIOTICLPMSYNDPOFGTSCITGKGENSTDYLPOLKMTVVKLISHREC 190
131 GRCAQPSRTIOTICLPMSYNDPOFGTSCITGKGENSTDYLPOLKMTVVKLISHREC 209
210 SRCQSSSVRTVCLPPADLPDPWTECELSGKHEALSPFYSERLKEAHLVLPSSRC 269
191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQCGDGGPLVCSLOGRMTLTGIVSWGRG 244
270 TSOHLNRTVTDNMLCAGDTRSGGPQANLHDAACQDGGPLVCLNDGRMTLVGLISWGLG 329
330 CQKXDPGVYTKVNTYLDNR 354

US-08-217-617A-1
Sequence 1, Application US/08217616
GENERAL INFORMATION:
APPLICANT: Martin, Ulrich
TITLE OF INVENTION: METHOD FOR TREATING THROMBOEMBOLIC CONDITIONS USING THROMBOLYTICALLY ACTIVE PROTEINS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS: 28
CITY: Lilly Lilly and Company
STATE: Indianapolis
COUNTRY: IN
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
OPERATING SYSTEM: Macintosh
CURRENT APPLICATION NUMBER: US/08/427,640
FILING DATE: 07/689,410
SEQUENCE CHARACTERISTICS: 6:
LENGTH: 355 amino acids
TOPOLOGY: linear
US-08-217-617A-1

Query Match
Best Local Similarity 38.7%; Score 583; DB 1; Length 355;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
13 CG-KTLRPRFKLIGGEFTTINQWFAAIYRHH-RGGSVTVYCGSLSPCWVISAHC 70
92 CGLRQYQSPQRIKGGFLADIASHPWQAIFAKHRSRSPGERFLCGGILISSCHWLSAHC 151
71 FIDYPKEDYVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
152 FOERFPPLHLLTVILGRSRTIOTICLPMSYNDPOFGTSCITGKGENSTDYLPOLKMTVVKLISHREC 190
131 GRCAQPSRTIOTICLPMSYNDPOFGTSCITGKGENSTDYLPOLKMTVVKLISHREC 209
210 SRCQSSSVRTVCLPPADLPDPWTECELSGKHEALSPFYSERLKEAHLVLPSSRC 269
191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQCGDGGPLVCSLOGRMTLTGIVSWGRG 244
270 TSOHLNRTVTDNMLCAGDTRSGGPQANLHDAACQDGGPLVCLNDGRMTLVGLISWGLG 329
330 CQKXDPGVYTKVNTYLDNR 354

US-09-880-503-5.ra1

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS: 1
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
OPERATING SYSTEM: PC-DOS
CURRENT APPLICATION NUMBER: US/08/217,617A
FILING DATE: 25-MARCH-1994
CLASSIFICATION: 424
NAME: Hanson, NO
REGISTRATION NUMBER: 5676947man D.
TELEPHONE: (212) 688-9200
TELEFAX: (212) 688-9200
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 355 amino acids
TOPOLOGY: linear
US-09-880-503-5.ra1

Query Match
Best Local Similarity 38.7%; Score 583; DB 1; Length 355;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
13 CG-KTLRPRFKLIGGEFTTINQWFAAIYRHH-RGGSVTVYCGSLSPCWVISAHC 70
92 CGLRQYQSPQRIKGGFLADIASHPWQAIFAKHRSRSPGERFLCGGILISSCHWLSAHC 151
71 FIDYPKEDYVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
152 FOERFPPLHLLTVILGRSRTIOTICLPMSYNDPOFGTSCITGKGENSTDYLPOLKMTVVKLISHREC 190
131 GRCAQPSRTIOTICLPMSYNDPOFGTSCITGKGENSTDYLPOLKMTVVKLISHREC 209
210 SRCQSSSVRTVCLPPADLPDPWTECELSGKHEALSPFYSERLKEAHLVLPSSRC 269
191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQCGDGGPLVCSLOGRMTLTGIVSWGRG 244
270 TSOHLNRTVTDNMLCAGDTRSGGPQANLHDAACQDGGPLVCLNDGRMTLVGLISWGLG 329
330 CQKXDPGVYTKVNTYLDNR 354

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2004, 12:58:25 ; Search time 21 Seconds

(without alignments)

1263.931 Million cell updates/sec

Title: US-09-880-503-5

Perfect score: 1508

Sequence: 1 KPSSPEELKFQCGQKTLRP.....VSHFLPWIRSHTEKENGAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1505	99.8	431	1 UKHU	u-plasminogen acti
2	1434.5	95.1	433	1 UKBAY	u-plasminogen acti
3	1255	83.2	442	1 UKPG	u-plasminogen acti
4	1188	78.8	433	1 JN0560	u-plasminogen acti
5	1119.5	74.2	432	1 S18932	u-plasminogen acti
6	1112.5	73.8	433	1 UKMS	u-plasminogen acti
7	688.5	45.7	434	1 A35005	u-plasminogen acti
8	583	38.7	582	1 UKHUT	u-plasminogen acti
9	570.5	37.8	559	1 A35029	t-plasminogen acti
10	569.5	37.8	559	1 A29941	t-plasminogen acti
11	547	36.3	431	2 JS0599	t-plasminogen acti
12	546	36.2	394	2 JS0600	t-plasminogen acti
13	546	36.2	477	2 JS0598	t-plasminogen acti
14	545	36.1	477	2 A34369	t-plasminogen acti
15	543	36.0	477	2 JS0597	t-plasminogen acti
16	495	32.8	615	1 KFHU12	coagulation factor
17	493.5	32.7	655	1 A46688	hepatocyte growth
18	485.5	32.2	603	2 S28941	coagulation factor
19	453	30.0	593	2 S45281	coagulation factor
20	425.5	28.2	761	2 JC5759	brain-specific ser
21	425	28.2	558	2 JC5878	plasma hyaluronan-
22	417	27.7	560	1 JC4795	proctasin (EC 3.4.
23	408.5	27.1	638	1 A57014	plasma kallikrein
24	407	27.0	638	1 KQMSPL	plasma kallikrein
25	402	26.7	638	1 KQHUP	hepsin (EC 3.4.21
26	396.5	26.3	417	1 S00845	trypsin (EC 3.4.21
27	395.5	26.2	248	2 S55066	trypsin (EC 3.4.21
28	394.5	26.2	229	1 TRROTR	plasmin (EC 3.4.21
29	392.5	26.0	460	2 B61545	

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N:Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen
N:Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form
C:Species: Homo sapiens (man)
C>Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C:Accession: A00931; 152209; J0102; A37561; 138102; S6578; A37562; A37563; A37564; A
R:Riccio, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Biasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A>Title: The human urokinase-plasminogen activator gene and its promoter.
A:Reference number: A00931; MUID:85215647; PMID:2987867
A:Accession: A00931
A:Molecule type: DNA
A:Cross-references: GB:X02419; NID:G37601; PIDN:CAA36268.1; PID:G1834524
A>Note: the authors translated the codon ATG for residue 214 as Ile
R:Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985
A>Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin
A:Reference number: 152209; MUID:86050639; PMID:19333505
A:Accession: 152209
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 145-161 <NAG1>
A:Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama.
Gene 36, 183-188, 1985
A>Title: Molecular cloning of cDNA coding for human preprourokinase.
A:Reference number: J0102; MUID:86056954; PMID:2415429
A:Accession: J0102
A:Molecule type: mRNA
A:Residues: 1-213, 'I', 215-431 <NAG2>
A:Cross-references: GB:X03226; NID:G340155; PIDN:AA97138.1; PID:G340158; GB:D00244; N
R:Verde, P.; Scoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984
A>Title: Identification and primary sequence of an unspliced human urokinase poly(A) + 1
A:Reference number: A37561; MUID:84272706; PMID:6589620
A:Accession: A37561
A:Molecule type: mRNA
A:Residues: 66-431 <VER>
A:Cross-references: GB:D00244; NID:G220138
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else
DNA 4, 139-146, 1985
A>Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr
A:Reference number: 138102; MUID:85203359; PMID:3888571
A:Accession: 138102
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A:Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA36535.1; PID:G35298

ALIGNMENTS

30	390.5	25.9	247	2	S13813	trypsin (EC 3.4.21
31	390	25.9	263	2	A11195	chymotrypsin (EC 3
32	386.5	25.6	416	1	S33777	hepsin (EC 3.4.21.
33	385	25.5	638	1	KQRTPL	plasma kallikrein
34	383.5	25.4	790	1	PLPG	plasmin (EC 3.4.21
35	382	25.3	263	1	KVRTB	chymotrypsin (EC 3
36	380	25.2	269	2	A26823	pancreatic elastase
37	379.5	25.2	304	2	S3496	trypsin (EC 3.4.21
38	378.5	25.1	264	2	S13136	chymotrypsin-like
39	378	25.1	259	2	S138363	trypsin (EC 3.4.21
40	377	25.0	263	2	A31299	chymotrypsin (EC 3
41	376.5	25.0	247	1	A25852	trypsin (EC 3.4.21
42	375	24.9	812	1	PLMS	plasmin (EC 3.4.21
43	374	24.8	271	1	ELRT2	pancreatic elastase
44	373.5	24.8	247	1	TRDG	trypsin (EC 3.4.21
45	373.5	24.8	261	2	A25606	tissue kallikrein

C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease

F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted <SIG>
F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <MI>
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <MI>
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <MI>
F;179-419/Domain: trypsin homology <TRY>
F;31-39,33-51,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-361,372-431/Product: urokinase-type plasminogen activator (covalent) #status predicted
F;36/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F;224-275,376/Active site: His, Asp, Ser #status experimental
F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 99.8%; Score 1505; DB 1; Length 431;
Best Local Similarity 99.6%; Pred. No. 3.9e-123;
Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTYVCGSLIS 60
DB 156 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTYVCGSLMS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCIEITGFGKENSDDIYLPOLKMT 180
DB 276 IALLKIRSKGRCAQPSRTTQTICLPSMYNDPQGTSCIEITGFGKENSDDIYLPOLKMT 335

QY 181 VKKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCOGDSGGPLVCSLQGRMTLTGIVS 240
DB 336 VKKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCOGDSGGPLVCSLQGRMTLTGIVS 395

QY 241 WGRGCAIKDPGVYTRVSHFLPWIRSHRSTKEENGIAL 276
DB 396 WGRGCAIKDPGVYTRVSHFLPWIRSHRSTKEENGIAL 431

RESULT 2

UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #test_change 18-Jun-1999
C;Accession: S14687; S03651
R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W.
Nucleic Acids Res, 18, 3411, 1990
A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
A;Reference number: S14687; MUID:90287734; PMID:2113276
A;Accession: S14687
A;Molecule type: mRNA
A;Residues: 1-433 <AUY>
A;Cross-references: EMBL:X51935; NID:938130; PIDN:CAA36200.1; PID:938131
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C;Keywords: Glycoprotein; heterodimer; hydrolase; kringle; serine protease
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-176/Product: plasminogen activator chain A #status predicted <ACH>
F;30-61/Domain: EGF homology <EGF>
F;69-150/Domain: kringle homology <KRG>
F;178-433/Product: plasminogen activator chain B #status predicted <BCH>
F;178-421/Domain: trypsin homology <TRY>
F;167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted
F;223,274,378/Active site: His, Asp, Ser #status predicted
F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 95.1%; Score 1434.5; DB 1; Length 433;
Best Local Similarity 92.8%; Pred. No. 5.3e-117;
Matches 259; Conservative 13; Mismatches 4; Indels 3; Gaps 1;

QY 1 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTYVCGSLIS 60
DB 155 KPSSPPEELKFCQCKQLTRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTYVCGSLIS 214

QY 61 PCWISATHCFIDYPPKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 120
 Db 215 PCWVSATHCFINYPKEDYIVYLGSRSLNSNTGEMKFEVENLILHEDYSADTLAHND 274
 QY 121 IALLKIRSKGRCACQPSRTIOTICLPMSYNDPQ---FGTSCEITGFGKENSVDLYPEQL 177
 Db 275 IALLKIRSKGRCACQPSRTIOTICLPMSYNDPDPFGTSCEITGFGKENSVDLYPEQL 334
 QY 178 KMTVKVLSHRECCQPHYGVSEVTTKMLCAADPOWKTDSQCGSGGGLVCSLQGRMTLTG 237
 Db 335 KMTVKVLSHRECCQPHYGVSEVTTKMLCAADPOWKTDSQCGSGGGLVCSLQGRMTLTG 394
 QY 238 IVSGRGKALXDKPGVYTRVSHFLPWIRSHTKENGIAL 276
 Db 395 IVSGRGKALXDKPGVYTRVSHFLPWIRSHTKENGIAL 433
 RESULT 3
 UKPG u-plasminogen activator (EC 3.4.21.73) precursor - pig
 N/Alternate names: uPA
 C/Species: Sus scrofa domestica (domestic pig)
 C/Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
 C/Accession: A00932
 F/Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.
 A/Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
 A/Reference number: A00932; MUID:85087954; PMID:6096832
 A/Accession: A00932
 A/Molecule type: DNA
 A/Residues: 1-240, 'H', 242-442 <NAG1>
 A/Experimental source: kidney cell line LLC-PK1
 R/Nagamine, Y.
 submitted to the Protein Sequence Database, December 1986
 A/Reference number: A37566
 A/Contents: annotation; Correction to residue 241
 C/Genetics:
 A/Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
 A/Suprafamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F/1-20/DNA: signal sequence #status predicted <SIG>
 F/21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F/33-64/DNA: EGF homology <EGF>
 F/72-153/DNA: kringle homology <KRG>
 F/190-430/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F/152/Binding site: Carbohydrate (Asn)
 F/179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
 F/235, 286, 387/Active site: His, Asp, Ser #status predicted
 Query Match 83.2%; Score 1255; DB 1; Length 442;
 Best Local Similarity 82.1%; Pred. No. 2.2e-101;
 Matches 225; Conservative 23; Mismatches 26; Indels 0; Gaps 0;
 QY 2 PSSPPELKFQCGQKTLRPFKIIGESTTTIENQWFAIYRHRGSGVTVYCGSLISP 61
 Db 168 PFTPEKVEFCQCKALRPFKIIGKSTIENQWFAIYRHRGSGVTVYCGSLISP 227
 QY 62 CWVISATHCFIDYPPKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 121
 Db 228 CWVISATHCFINYPKEDYIVYLGSRSLNSNTGEMKFEVENLILHEDYSADTLAHND 287
 QY 122 ALLKIRSKGRCACQPSRTIOTICLPMSYNDPQGTSCETGFGKENSVDLYPEQLKMTV 181
 Db 288 ALLKIRSKGRCACQPSRTIOTICLPMSYNDPQGTSCETGFGKENSVDLYPEQLKMTV 347
 QY 182 VKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSQCGSGGGLVCSLQGRMTLTGIVSW 241
 Db 348 VKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSQCGSGGGLVCSLQGRMTLTGIVSW 407
 QY 242 GRGKALXDKPGVYTRVSHFLPWIRSHTKENGIAL 275
 Db 242 GRGKALXDKPGVYTRVSHFLPWIRSHTKENGIAL 275

Db 408 GRECAMKDPGVYTRVSRFLTWIHTHVGGENGIA 441
 RESULT 4
 JN0560 u-plasminogen activator (EC 3.4.21.73) precursor - bovine
 N/Alternate names: uPA
 C/Species: Bos primigenius taurus (cattle)
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C/Accession: JN0560
 R/Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
 Gene 125, 177-183, 1993
 A/Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indi
 A/Reference number: JN0560; MUID:93216119; PMID:8385052
 A/Accession: JN0560
 A/Molecule type: mRNA
 A/Residues: 1-433 <KRA>
 A/Cross-references: GB:L03546; NID:gl3800; PID:AAAS1419.1; PID:gl3801
 C/Suprafamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
 C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F/1-20/DNA: signal sequence #status predicted <SIG>
 F/21-179/Product: plasminogen activator chain A #status predicted <MA1>
 F/21-179/Product: plasminogen activator chain A #status predicted <ACH>
 F/33-64/DNA: EGF homology <EGF>
 F/72-153/DNA: kringle homology <KRG>
 F/181-433/Product: plasminogen activator chain B #status predicted <MA2>
 F/181-421/DNA: trypsin homology <TRY>
 F/170-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
 F/226, 277, 378/Active site: His, Asp, Ser #status predicted
 Query Match 78.8%; Score 1188; DB 1; Length 433;
 Best Local Similarity 76.4%; Pred. No. 1.5e-95;
 Matches 210; Conservative 32; Mismatches 33; Indels 0; Gaps 0;
 QY 2 PSSPPELKFQCGQKTLRPFKIIGESTTTIENQWFAIYRHRGSGVTVYCGSLISP 61
 Db 159 PSSPPELKFQCGQKTLRPFKIIGESTTTIENQWFAIYRHRGSGVTVYCGSLISP 218
 QY 62 CWVISATHCFIDYPPKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 121
 Db 219 CWVISATHCFIDYPPKEDYIVYLGSRSLNSNTGEMKFEVENLILHEDYSADTLAHND 278
 QY 122 ALLKIRSKGRCACQPSRTIOTICLPMSYNDPQGTSCETGFGKENSVDLYPEQLKMTV 181
 Db 279 ALLKIRSKGRCACQPSRTIOTICLPMSYNDPQGTSCETGFGKENSVDLYPEQLKMTV 338
 QY 182 VKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSQCGSGGGLVCSLQGRMTLTGIVSW 241
 Db 339 VSLVSHRECCQPHYGVSEVTTKMLCAADPOWKTDSQCGSGGGLVCSLQGRMTLTGIVSW 398
 QY 242 GRGKALXDKPGVYTRVSHFLPWIRSHTKENGIAL 276
 Db 399 GRGKALXDKPGVYTRVSHFLPWIRSHTKENGIAL 433
 RESULT 5
 S18932 u-plasminogen activator (EC 3.4.21.73) precursor - rat
 N/Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activato
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
 C/Accession: S24604; I53472; S18932
 R/Kabani, S.A.
 submitted to the EMBL Data Library, April 1992
 A/Reference number: S24604
 A/Accession: S24604
 A/Molecule type: mRNA
 A/Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
 A/Cross-references: EMBL:X65651; NID:S57456; PID:CAA46601.1; PID:G57457
 A/Experimental source: tissue kidney
 R/Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefferd, R.F.
 Cancer Res. 52, 2489-2496, 1992
 A/Title: Transcriptional and posttranscriptional activation of urokinase plasminogen a

A:Reference number: 160186; MUID:92233409; PMID:1568219
A:Accession: 160186
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-432 <RES>
A:Cross-references: EMBL:X63434; NID:957465; PIDN:CAA45028.1; PID:957466
A:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
A:Reference number: 153472; MUID:92339549; PMID:1321734
A:Accession: 153472
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 31-62 <RES>
A:Cross-references: EMBL:X66907; NID:9396200; PIDN:CAA47356.1; PID:9398279
C:Genetics: uPA
A:Gene: uPA
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; trypsin
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:31-62/Domain: EGF homology <EGF>
F:70-151/Domain: kringle homology <KRG>
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:179-420/Domain: trypsin homology <TRY>
F:168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted
F:225,276,377/Active site: His, Asp, Ser #status predicted
Query Match 74.2%; Score 1119.5; DB 1; Length 432;
Best Local Similarity 71.4%; Pred. No. 1.3e-89;
Matches 197; Conservative 39; Mismatches 39; Indels 1; Gaps 1;
Qy 1 KPSSPEELKFCQGOKTLPRPKIIGGFTTINQWPAALYRRHRGGS-VTVVCGGSLI 59
Db 156 KPSSTVDQGGFCQGGKALPRFKIVGGFTVVENQWPAALYLNKGGSPSPKCGGSLI 215
Qy 60 SPQWISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHNN 119
Db 216 SPQWASATHCFVNPQKKEEVVYVYLGQSKNSYNGEMKFEVEQLILHEDSDTLAHNN 275
Qy 120 DIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCITGFGKNSDYLYPEQLKM 179
Db 276 DIALLKIRTSQCAQPSRTIQTICLPRFGDAPFGSCITGFGQESATDYLYPKDLKM 335
Qy 180 TVVKLIHRECOQPHYGSEVTKMLCAADPWKTDSCGDSGGPLVCSLQGRMTLTGIV 239
Db 336 SVVLIISHFQCKOPHYGSEINVKMLCAADPWKTDSCGDSGGPLICNIDGRPTLSGIV 395
Qy 240 SWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 275
Db 396 SWGGCAEKNGKPGVYTRVSHFLPWIRSHTKKEENGLA 431
RESULT 6
UKMS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C:Accession: A29420; A24615
R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
Biochemistry 26, 8270-8279, 1987
A:Title: The murine urokinase-type plasminogen activator gene.
A:Reference number: A29420; MUID:88163489; PMID:2831940
A:Accession: A29420
A:Molecule type: DNA
A:Residues: 1-433 <RES>
A:Cross-references: GB:M17922; NID:9202286; PIDN:AAA40539.1; PID:9202297
R:Belin, D.; Vassalli, J.D.; Compeigne, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1985
A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
A:Reference number: A24615; MUID:85179474; PMID:2985383
A:Accession: A24615

A:Molecule type: mRNA
A:Residues: 1-433 <BEL>
A:Cross-references: GB:X02389; NID:955127; PIDN:CAA26231.1; PID:955128
C:Genetics: uPA
A:Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:32-63/Domain: EGF homology <EGF>
F:71-152/Domain: kringle homology <KRG>
F:180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:180-421/Domain: trypsin homology <TRY>
F:169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted
F:226,277,378/Active site: His, Asp, Ser #status predicted
Query Match 73.8%; Score 1112.5; DB 1; Length 433;
Best Local Similarity 71.0%; Pred. No. 5.4e-89;
Matches 196; Conservative 39; Mismatches 40; Indels 1; Gaps 1;
Qy 1 KPSSPEELKFCQGOKTLPRPKIIGGFTTINQWPAALYRRHRGGS-VTVVCGGSLI 59
Db 157 KPSSSVDDQGGFCQGGKALPRFKIVGGFTVVENQWPAALYLNKGGSPSPKCGGSLI 216
Qy 60 SPQWISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHNN 119
Db 217 SPQWASATHCFIQLPKENYVYVYLGQSKNSYNGEMKFEVEQLILHEVYREDSLAYHN 276
Qy 120 DIALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCITGFGKNSDYLYPEQLKM 179
Db 277 DIALLKIRTSQCAQPSRTIQTICLPRFTDAPFGSCITGFGKESDYLYPKNLKM 336
Qy 180 TVVKLIHRECOQPHYGSEVTKMLCAADPWKTDSCGDSGGPLVCSLQGRMTLTGIV 239
Db 337 SVVLIISHFQCKOPHYGSEINVKMLCAADPWKTDSCGDSGGPLICNIEGRPTLSGIV 396
Qy 240 SWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 275
Db 397 SWRGCAEKNGKPGVYTRVSHFLPWIRSHTKKEENGLA 432
RESULT 7
A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N:Alternate names: uPA
C:Species: Gallus gallus (chicken)
C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C:Accession: A35005
R:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A:Title: The chicken urokinase-type plasminogen activator gene.
A:Reference number: A35005; MUID:90110185; PMID:2295632
A:Accession: A35005
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <LES>
A:Cross-references: GB:J05187; NID:9212858; PIDN:AAA49131.1; PID:9212859
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:40-71/Domain: EGF homology <EGF>
F:179-158/Domain: kringle homology <KRG>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCI>
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted
Query Match 45.7%; Score 688.5; DB 1; Length 434;
Best Local Similarity 47.4%; Pred. No. 4.1e-52;
Matches 128; Conservative 48; Mismatches 89; Indels 5; Gaps 2;
Qy 2 PSSPEELKFCQGOKTLPRPKIIGGFTTINQWPAALYRRHRGGS-VTVVCGGSLISP 61

Dd 151 PSTIEKERTCQRFSEKFKLVGSGQAEVETQPIAGIFQNM-GTQPLCGGSLIP 209
Qy 62 CWSIATHCFID----YPKEDYIVVLGRSLNSNTQGMKFEVNLILHKQYSADTLAH 117
Dd 210 CWVLTAAHCFYNETKKQPKNSVKYKFLGKSLNTNDEHEQVFWVDIISHPDFTDHTGNN 269
Qy 118 HNDIALLKRSKGRCAQPSRTTCTILPSMYNDPQFGTSCETITGFGKNSDYLYPEOL 177
Dd 270 DNDIALIRITAGCCAVESNYVTRTVCLPEKNLUNLDVNTWCEIAGTKQNSYDIYQAQL 329
Qy 178 KMTVWKLISHRECOQPHYGYSEVTTKMLCAADPQWTKDSQCGSGPLVCSLQGRMTLIG 237
Dd 330 MSATVNLISQDDCKNKYDSTRVTDNMVCAQDPLWETDACKGDSGGMVCEHNGRMTLYG 369
Qy 238 IVSWGRCALKDKPGVYVTRVSHPLPIRSH 267
Dd 390 IVSWGDCGCKNKGVYVTRVRLNWDISN 419
RESULT 8
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N:Alternate names: t-PA; tissue plasminogen activator
C:Species: Homo sapiens (man)
C:Date: 14-Nov-1983 #sequence, revision 14-Nov-1983 #text, change 08-Dec-2000
C:Accession: A94004; A23529; J070562; A93293; S02125; A91343; A93951; A91322; A54645; I60
R:NY, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A:Title: The structure of the human tissue-type plasminogen activator gene: correlation
A:Reference number: A94004; MUID:84298137; PMID:6089198
A:Accession: A94004
A:Molecule type: DNA
A:Residues: 1-562 <NYT>
A:Cross-references: GB:L00141
A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
R:Pierneer Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A:Title: The human tissue plasminogen activator gene.
A:Reference number: A23529; MUID:86196143; PMID:3009482
A:Accession: A23529
A:Molecule type: DNA
A:Residues: 1-562 <DEG>
A:Cross-references: GB:K03021; NID:g339817; PIDN:AAA98809.1; PID:g339818
R:Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A:Title: Purification and characterization of tissue plasminogen activator secreted by H
R:Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1993
A:Reference number: J070562; MUID:91291340; PMID:1368681
A:Accession: J070562
A:Molecule type: mRNA
A:Residues: 31-562 <ITA>
A:Cross-references: DDBJ:D01096; NID:g220128; PIDN:BAA00881.1; PID:g441174
A:Experimental source: embryonic lung fibroblast IMR-90 cells
A:Note: part of this sequence, including the amino end of the mature protein, was confir
R:Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1993
A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A:Reference number: A93293; MUID:83115262; PMID:6337343
A:Accession: A93293
A:Molecule type: mRNA
A:Residues: 1-562 <PEN>
A:Cross-references: GB:L00141
A:Experimental source: melanoma cells
R:Sasaki, H.; Saito, Y.; Hayaishi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A:Reference number: S02125; MUID:88262579; PMID:3133640
A:Accession: S02125
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-562 <SAS>
A:Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244
A:Experimental source: fetal lung cells

R:Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; P
FEBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminoge
A:Reference number: A91343; MUID:85285620; PMID:3896853
A:Accession: A91343
A:Molecule type: mRNA
A:Residues: 1-38, 'G', '86-433', 'E', '435-562' <KAG>
A:Experimental source: Detroit 562 cells; ATCC 138
R:Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen acti
A:Reference number: A93951; MUID:83169656; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
A:Experimental source: melanoma cells
R:Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived
differences.
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation: melanoma cells, partial sequence of residues 36-562, active anc
R:Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator:
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322
A:Molecule type: protein
A:Residues: 33-45, 311-320 <POH>
A:Experimental source: uterus
A:Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R:Van Zonneveld, A.J.; Veerman, H.; Pannexoe, H.
J. Biol. Chem. 261, 14214-14218, 1986
A:Reference number: A37567; MUID:87033611; PMID:3021732
A:Contents: annotation: fibrin binding site
R:Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Eng
EMBO J. 5, 3525-3530, 1986
A:Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen
A:Reference number: A37568; MUID:87161761; PMID:3030730
A:Contents: annotation: fibrin binding site
R:Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A:Title: Isolation, identification and pharmacokinetic properties of human tissue-type
A:Reference number: A60902; MUID:89044681; PMID:3142086
A:Contents: annotation: novel forms of expressed recombinant t-PA
R:Haris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.
Mol. Biol. Med. 3, 279-292, 1985
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its exp
A:Reference number: A54645; MUID:86284200; PMID:3090401
A:Accession: A54645
A:Molecule type: mRNA
A:Residues: 1-562 <HAR>
A:Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032
A:Note: parts of this sequence were confirmed by peptide sequencing
R:Reddy, V.B.; Garrazone, A.J.; Sasaki, H.; Wei, C.
DNA 6, 461-472, 1987
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A:Reference number: I60110; MUID:88054470; PMID:2824147
A:Accession: I60110
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-562 <RES>
A:Cross-references: GB:M18192; NID:g340176; PIDN:AAA36800.1; PID:g340177
R:Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11231-11230, 1985
A:Title: Isolation and characterization of the human tissue-type plasminogen activator
A:Reference number: I55232; MUID:85289338; PMID:3161893
A:Accession: I55232
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-36 <RE2>
A:Cross-references: GB:M11890; NID:g339837; PIDN:AAA61213.1; PID:g339839
C:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single

F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <IFI>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000,1001,1002,1003,1004,1005,1006,1007,1008,1009,1010,1011,1012,1013,1014,1015,1016,1017,1018,1019,1020,1021,1022,1023,1024,1025,1026,1027,1028,1029,1030,1031,1032,1033,1034,1035,1036,1037,1038,1039,1040,1041,1042,1043,1044,1045,1046,1047,1048,1049,1050,1051,1052,1053,1054,1055,1056,1057,1058,1059,1060,1061,1062,1063,1064,1065,1066,1067,1068,1069,1070,1071,1072,1073,1074,1075,1076,1077,1078,1079,1080,1081,1082,1083,1084,1085,1086,1087,1088,1089,1090,1091,1092,1093,1094,1095,1096,1097,1098,1099,1100,1101,1102,1103,1104,1105,1106,1107,1108,1109,1110,1111,1112,1113,1114,1115,1116,1117,1118,1119,1120,1121,1122,1123,1124,1125,1126,1127,1128,1129,1130,1131,1132,1133,1134,1135,1136,1137,1138,1139,1140,1141,1142,1143,1144,1145,1146,1147,1148,1149,1150,1151,1152,1153,1154,1155,1156,1157,1158,1159,1160,1161,1162,1163,1164,1165,1166,1167,1168,1169,1170,1171,1172,1173,1174,1175,1176,1177,1178,1179,1180,1181,1182,1183,1184,1185,1186,1187,1188,1189,1190,1191,1192,1193,1194,1195,1196,1197,1198,1199,1200,1201,1202,1203,1204,1205,1206,1207,1208,1209,1210,1211,1212,1213,1214,1215,1216,1217,1218,1219,1220,1221,1222,1223,1224,1225,1226,1227,1228,1229,1230,1231,1232,1233,1234,1235,1236,1237,1238,1239,1240,1241,1242,1243,1244,1245,1246,1247,1248,1249,1250,1251,1252,1253,1254,1255,1256,1257,1258,1259,1260,1261,1262,1263,1264,1265,1266,1267,1268,1269,1270,1271,1272,1273,1274,1275,1276,1277,1278,1279,1280,1281,1282,1283,1284,1285,1286,1287,1288,1289,1290,1291,1292,1293,1294,1295,1296,1297,1298,1299,1300,1301,1302,1303,1304,1305,1306,1307,1308,1309,1310,1311,1312,1313,1314,1315,1316,1317,1318,1319,1320,1321,1322,1323,1324,1325,1326,1327,1328,1329,1330,1331,1332,1333,1334,1335,1336,1337,1338,1339,1340,1341,1342,1343,1344,1345,1346,1347,1348,1349,1350,1351,1352,1353,1354,1355,1356,1357,1358,1359,1360,1361,1362,1363,1364,1365,1366,1367,1368,1369,1370,1371,1372,1373,1374,1375,1376,1377,1378,1379,1380,1381,1382,1383,1384,1385,1386,1387,1388,1389,1390,1391,1392,1393,1394,1395,1396,1397,1398,1399,1400,1401,1402,1403,1404,1405,1406,1407,1408,1409,1410,1411,1412,1413,1414,1415,1416,1417,1418,1419,1420,1421,1422,1423,1424,1425,1426,1427,1428,1429,1430,1431,1432,1433,1434,1435,1436,1437,1438,1439,1440,1441,1442,1443,1444,1445,1446,1447,1448,1449,1450,1451,1452,1453,1454,1455,1456,1457,1458,1459,1460,1461,1462,1463,1464,1465,1466,1467,1468,1469,1470,1471,1472,1473,1474,1475,1476,1477,1478,1479,1480,1481,1482,1483,1484,1485,1486,1487,1488,1489,1490,1491,1492,1493,1494,1495,1496,1497,1498,1499,1500,1501,1502,1503,1504,1505,1506,1507,1508,1509,1510,1511,1512,1513,1514,1515,1516,1517,1518,1519,1520,1521,1522,1523,1524,1525,1526,1527,1528,1529,1530,1531,1532,1533,1534,1535,1536,1537,1538,1539,1540,1541,1542,1543,1544,1545,1546,1547,1548,1549,1550,1551,1552,1553,1554,1555,1556,1557,1558,1559,1560,1561,1562,1563,1564,1565,1566,1567,1568,1569,1570,1571,1572,1573,1574,1575,1576,1577,1578,1579,1580,1581,1582,1583,1584,1585,1586,1587,1588,1589,1590,1591,1592,1593,1594,1595,1596,1597,1598,1599,1600,1601,1602,1603,1604,1605,1606,1607,1608,1609,1610,1611,1612,1613,1614,1615,1616,1617,1618,1619,1620,1621,1622,1623,1624,1625,1626,1627,1628,1629,1630,1631,1632,1633,1634,1635,1636,1637,1638,1639,1640,1641,1642,1643,1644,1645,1646,1647,1648,1649,1650,1651,1652,1653,1654,1655,1656,1657,1658,1659,1660,1661,1662,1663,1664,1665,1666,1667,1668,1669,1670,1671,1672,1673,1674,1675,1676,1677,1678,1679,1680,1681,1682,1683,1684,1685,1686,1687,1688,1689,1690,1691,1692,1693,1694,1695,1696,1697,1698,1699,1700,1701,1702,1703,1704,1705,1706,1707,1708,1709,1710,1711,1712,1713,1714,1715,1716,1717,1718,1719,1720,1721,1722,1723,1724,1725,1726,1727,1728,1729,1730,1731,1732,1733,1734,1735,1736,1737,1738,1739,1740,1741,1742,1743,1744,1745,1746,1747,1748,1749,1750,1751,1752,1753,1754,1755,1756,1757,1758,1759,1760,1761,1762,1763,1764,1765,1766,1767,1768,1769,1770,1771,1772,1773,1774,1775,1776,1777,1778,1779,1780,1781,1782,1783,1784,1785,1786,1787,1788,1789,1790,1791,1792,1793,1794,1795,1796,1797,1798,1799,1800,1801,1802,1803,1804,1805,1806,1807,1808,1809,1810,1811,1812,1813,1814,1815,1816,1817,1818,1819,1820,1821,1822,1823,1824,1825,1826,1827,1828,1829,1830,1831,1832,1833,1834,1835,1836,1837,1838,1839,1840,1841,1842,1843,1844,1845,1846,1847,1848,1849,1850,1851,1852,1853,1854,1855,1856,1857,1858,1859,1860,1861,1862,1863,1864,1865,1866,1867,1868,1869,1870,1871,1872,1873,1874,1875,1876,1877,1878,1879,1880,1881,1882,1883,1884,1885,1886,1887,1888,1889,1890,1891,1892,1893,1894,1895,1896,1897,1898,1899,1900,1901,1902,1903,1904,1905,1906,1907,1908,1909,1910,1911,1912,1913,1914,1915,1916,1917,1918,1919,1920,1921,1922,1923,1924,1925,1926,1927,1928,1929,1930,1931,1932,1933,1934,1935,1936,1937,1938,1939,1940,1941,1942,1943,1944,1945,1946,1947,1948,1949,1950,1951,1952,1953,1954,1955,1956,1957,1958,1959,1960,1961,1962,1963,1964,1965,1966,1967,1968,1969,1970,1971,1972,1973,1974,1975,1976,1977,1978,1979,1980,1981,1982,1983,1984,1985,1986,1987,1988,1989,1990,1991,1992,1993,1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021,2022,2023,2024,2025,2026,2027,2028,2029,2030,2031,2032,2033,2034,2035,2036,2037,2038,2039,2040,2041,2042,2043,2044,2045,2046,2047,2048,2049,2050,2051,2052,2053,2054,2055,2056,2057,2058,2059,2060,2061,2062,2063,2064,2065,2066,2067,2068,2069,2070,2071,2072,2073,2074,2075,2076,2077,2078,2079,2080,2081,2082,2083,2084,2085,2086,2087,2088,2089,2090,2091,2092,2093,2094,2095,2096,2097,2098,2099,2100,2101,2102,2103,2104,2105,2106,2107,2108,2109,2110,2111,2112,2113,2114,2115,2116,2117,2118,2119,2120,2121,2122,2123,2124,2125,2126,2127,2128,2129,2130,2131,2132,2133,2134,2135,2136,2137,2138,2139,2140,2141,2142,2143,2144,2145,2146,2147,2148,2149,2150,2151,2152,2153,2154,2155,2156,2157,2158,2159,2160,2161,2162,2163,2164,2165,2166,2167,2168,2169,2170,2171,2172,2173,2174,2175,2176,2177,2178,2179,2180,2181,2182,2183,2184,2185,2186,2187,2188,2189,2190,2191,2192,2193,2194,2195,2196,2197,2198,2199,2200,2201,2202,2203,2204,2205,2206,2207,2208,2209,2210,2211,2212,2213,2214,2215,2216,2217,2218,2219,2220,2221,2222,2223,2224,2225,2226,2227,2228,2229,2230,2231,2232,2233,2234,2235,2236,2237,2238,2239,2240,2241,2242,2243,2244,2245,2246,2247,2248,2249,2250,2251,2252,2253,2254,2255,2256,2257,2258,2259,2260,2261,2262,2263,2264,2265,2266,2267,2268,2269,2270,2271,2272,2273,2274,2275,2276,2277,2278,2279,2280,2281,2282,2283,2284,2285,2286,2287,2288,2289,2290,2291,2292,2293,2294,2295,2296,2297,2298,2299,2300,2301,2302,2303,2304,2305,2306,2307,2308,2309,2310,2311,2312,2313,2314,2315,2316,2317,2318,2319,2320,2321,2322,2323,2324,2325,2326,2327,2328,2329,2330,2331,2332,2333,2334,2335,2336,2337,2338,2339,2340,2341,2342,2343,2344,2345,2346,2347,2348,2349,2350,2351,2352,2353,2354,2355,2356,2357,2358,2359,2360,2361,2362,2363,2364,2365,2366,2367,2368,2369,2370,2371,2372,2373,2374,2375,2376,2377,2378,2379,2380,2381,2382,2383,2384,2385,2386,2387,2388,2389,2390,2391,2392,2393,2394,2395,2396,2397,2398,2399,2400,2401,2402,2403,2404,2405,2406,2407,2408,2409,2410,2411,2412,2413,2414,2415,2416,2417,2418,2419,2420,2421,2422,2423,2424,2425,2426,2427,2428,2429,2430,2431,2432,2433,2434,2435,2436,2437,2438,2439,2440,2441,2442,2443,2444,2445,2446,2447,2448,2449,2450,2451,2452,2453,2454,2455,2456,2457,2458,2459,2460,2461,2462,2463,2464,2465,2466,2467,2468,2469,2470,2471,2472,2473,2474,2475,2476,2477,2478,2479,2480,2481,2482,2483,2484,2485,2486,2487,2488,2489,2490,2491,2492,2493,2494,2495,2496,2497,2498,2499,2500,2501,2502,2503,2504,2505,2506,2507,2508,2509,2510,2511,2512,2513,2514,2515,2516,2517,2518,2519,2520,2521,2522,2523,2524,2525,2526,2527,2528,2529,2530,2531,2532,2533,2534,2535,2536,2537,2538,2539,2540,2541,2542,2543,2544,2545,2546,2547,2548,2549,2550,2551,2552,2553,2554,2555,2556,2557,2558,2559,2560,2561,2562,2563,2564,2565,2566,2567,2568,2569,2570,2571,2572,2573,2574,2575,2576,2577,2578,2579,2580,2581,2582,2583,2584,2585,2586,2587,2588,2589,2590,2591,2592,2593,2594,2595,2596,2597,2598,2599,2600,2601,2602,2603,2604,2605,2606,2607,2608,2609,2610,2611,2612,2613,2614,2615,2616,2617,2618,2619,2620,2621,2622,2623,2624,2625,2626,2627,2628,2629,2630,2631,2632,2633,2634,2635,2636,2637,2638,2639,2640,2641,2642,2643,2644,2645,2646,2647,2648,2649,2650,2651,2652,2653,2654,2655,2656,2657,2658,2659,2660,2661,2662,2663,2664,2665,2666,2667,2668,2669,2670,2671,2672,2673,2674,2675,2676,2677,2678,2679,2680,2681,2682,2683,2684,2685,2686,2687,2688,2689,2690,2691,2692,2693,2694,2695,2696,2697,2698,2699,2700,2701,2702,2703,2704,2705,2706,2707,2708,2709,2710,2711,2712,2713,2714,2715,2716,2717,2718,2719,2720,2721,2722,2723,2724,2725,2726,2727,2728,2729,2730,2731,2732,2733,2734,2735,2736,2737,2738,2739,2740,2741,2742,2743,2744,2745,2746,2747,2748,2749,2750,2751,2752,2753,2754,2755,2756,2757,2758,2759,2760,2761,2762,2763,2764,2765,2766,2767,2768,2769,2770,2771,2772,2773,2774,2775,2776,2777,2778,2779,2780,2781,2782,2783,2784,2785,2786,2787,2788,2789,2790,2791,2792,2793,2794,2795,2796,2797,2798,2799,2800,2801,2802,2803,2804,2805,2806,2807,2808,2809,2810,2811,2812,2813,2814,2815,2816,2817,2818,2819,2820,2821,2822,2823,2824,2825,2826,2827,2828,2829,2830,2831,2832,2833,2834,2835,2836,2837,2838,2839,2840,2841,2842,2843,2844,2845,2846,2847,2848,2849,2850,2851,2852,2853,2854,2855,2856,2857,2858,2859,2860,2861,2862,2863,2864,2865,2866,2867,2868,2869,2870,2871,2872,2873,2874,2875,2876,2877,2878,2879,2880,2881,2882,2883,2884,2885,2886,2887,2888,2889,2890,2891,2892,2893,2894,2895,2896,2897,2898,2899,2900,2901,2902,2903,2904,2905,2906,2907,2908,2909,2910,29

RESULT 13

JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0598
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0598
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63988; NID:G166074; PID:AAA31593.1; PID:G166075
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringlie; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.2%; Score 546; DB 2; Length 477;
Best Local Similarity 42.9%; Pred. No. 1.4e-39;
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;

QY 4 SPPEELKFCQG-OKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGS-VTVVCGGSLISP 61
DB 205 SVPVCSKATGLRYKKEPQLHSTGGLFDITSHPWQAIFAQNRSSGERFLCGGILISS 264
QY 62 CWVISATHCFID-YPKXEDYIVYLGSRSLNNTQGMKFEVENILHKYSDTLAHND 120
DB 265 CWVLTAAHCFQERYPPQHLRVV-LGRTRYVRKPGKEEQTFVEKCIHVHEPDDDT--YNNND 321
QY 121 IALLKIRSKRGCRQPSRTTICTCLPSMYNDPQFGTSCEITGFGKENSITDLYLPEOLKMT 180
DB 322 IALLQKSGSPQCAQESDSVRAICLPEANLQLPDWTECELSYGGKHKSSFFYSQLEK 381
QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKT-----DSCQGDGSGPLVCSLQGRMT 234
DB 382 HVRLYPSSRCTSKFLPNKTVNNMLCAGDTRSGEIPNVHDAQCQDGGPLVCMNDNMT 441
QY 235 LTGIVSWGRGKALKDKPGVTVRVSHFLPWIRSHTK 269
DB 442 LLGIISWGVGCGEKDIPGVYTKVNYLGMWIRDNMR 476
RESULT 14
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C:Species: Megaderma lyra
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A34369
R:Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
J. Biol. Chem. 264, 17947-17952, 1989
A:Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A:Reference number: A34369; MUID:90036867; PMID:2509450
A:Accession: A34369
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <GAR>
A:Cross-references: GB:J05082; NID:G166080; PID:AAA31596.1; PID:G166081
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringlie; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>

F:37-477/Product: plasminogen activator #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: trypsin homology <TRY>
F:226-471/Domain: signal sequence #status predicted <SIG>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.1%; Score 545; DB 1; Length 477;
Best Local Similarity 42.9%; Pred. No. 1.4e-39;
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;

QY 4 SPPEELKFCQG-OKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGS-VTVVCGGSLISP 61
DB 205 SVPVCSKATGLRYKKEPQLHSTGGLFDITSHPWQAIFAQNRSSGERFLCGGILISS 264
QY 62 CWVISATHCFID-YPKXEDYIVYLGSRSLNNTQGMKFEVENILHKYSDTLAHND 120
DB 265 CWVLTAAHCFQERYPPQHLRVV-LGRTRYVRKPGKEEQTFVEKCIHVHEPDDDT--YNNND 321
QY 121 IALLKIRSKRGCRQPSRTTICTCLPSMYNDPQFGTSCEITGFGKENSITDLYLPEOLKMT 180
DB 322 IALLQKSGSPQCAQESDSVRAICLPEANLQLPDWTECELSYGGKHKSSFFYSQLEK 381
QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKT-----DSCQGDGSGPLVCSLQGRMT 234
DB 382 HVRLYPSSRCTSKFLPNKTVNNMLCAGDTRSGEIPNVHDAQCQDGGPLVCRNDNMT 441
QY 235 LTGIVSWGRGKALKDKPGVTVRVSHFLPWIRSHTK 269
DB 442 LLGIISWGVGCGEKDIPGVYTKVNYLGMWIRDNMR 476

RESULT 15

JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0597
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Do
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0597
A:Molecule type: mRNA
A:Residues: 1-477 <KRA>
A:Cross-references: GB:M63987; NID:G166070; PID:AAA31591.1; PID:G166071
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringlie; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <1FA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359
F:153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.0%; Score 543; DB 2; Length 477;
Best Local Similarity 42.2%; Pred. No. 2e-39;
Matches 117; Conservative 42; Mismatches 98; Indels 20; Gaps 6;

QY 4 SPPEELKFCQG-OKTLRPRFKIIGGEFTTIENQWPFAAIYRRHRGGS-VTVVCGGSLISP 61
DB 205 SVPVCSKATGLRYKKEPQLHSTGGLFDITSHPWQAIFAQNRSSGERFLCGGILISS 264
QY 62 CWVISATHCFIDYPKXEDYI-----VYLGSRSLNNTQGMKFEVENILHKYSDTLA 116
DB 265 CWVLTAAHCF-----QESYLPDLKXVLGRTYVRKPGEEQTFVKYKIVHKEFDDDT-- 317

QY 117 HNDIALKIRSGCAOPSRITQICLPMSVNDPQFCTSCITGFGKSTNDYLPQ 176
Db 318 YNDIALLKDSPOCAESVSRAICLPANLQLPDWTBELSGYKHKSSFFYSQ 377
QY 177 LKNTVVKLJSHRECQPHYVSEVTTMLCAADPQWKT-----DSCQDGGPLVCSLQ 230
Db 378 LKEGHVRLYPSRCAPKFLFNKTVTNMLCAGTRSGEYTPNVHDACQDGGPLVCMND 437
QY 231 GRMTLTGIVSWGRCALKDKPGVYTVSHFLPWIRSH 267
Db 438 NMHTLLGIISWVGCGEKDPVGYTTKVTNYLGMIRDN 474
RESULT 16
KFHU12
coagulation factor XIia (EC 3.4.21.38) precursor [validated] - human
N;Alternate names: Hageman factor (activated)
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 30-Jun-1991 #text_change 08-Dec-2000
C;Accession: A29411; A26814; A00930; A25191; A22248; A21037
R;Cool, D.E.; MacGillivray, R.T.A.
J. Biol. Chem. 262, 13662-13673, 1987
A;Title: Characterization of the human blood coagulation factor XII gene. Intron/exon 9
A;Reference number: A29411; MUID:88007593; PMID:2888762
A;Accession: A29411
A;Molecule type: DNA
A;Residues: 1-615 <COO>
A;Cross-references: GB:M17466; GB:J02807; NID:G180357; PIDN:AA59490.1; PID:G180357
R;Triodi, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.
Nucleic Acids Res. 14, 3146, 1986
A;Title: CDNA sequence coding for human coagulation factor XII (Hageman).
A;Reference number: A26814; MUID:86176794; PMID:3754331
A;Accession: A26814
A;Molecule type: mRNA
A;Residues: 4-615 <TRI>
A;Cross-references: GB:M31315; NID:G182291; PIDN:AAA70225.1; PID:G182292
R;Cool, D.E.; Eggeil, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R.T.
J. Biol. Chem. 260, 13666-13676, 1985
A;Title: Characterization of human blood coagulation factor XII cDNA. Prediction of the
A;Reference number: A00930; MUID:86033830; PMID:3877053
A;Accession: A00930
A;Molecule type: mRNA
A;Residues: 14-332, 'S', 334-615 <CO2>
A;Cross-references: GB:M11723; NID:G180358; PIDN:AAA51986.1; PID:G180359
R;Que, B.G.; Davie, E.W.
Biochemistry 25, 1525-1528, 1986
A;Title: Characterization of a cDNA coding for human factor XII (Hageman factor).
A;Reference number: A25191; MUID:86216049; PMID:3011063
A;Accession: A25191
A;Molecule type: mRNA
A;Residues: 146-378, 'G', 380-615 <QUB>
A;Cross-references: GB:M13147; NID:G180360; PIDN:AAA70224.1; PID:G180361
R;McMullen, B.A.; Fujikawa, K.
J. Biol. Chem. 260, 5328-5341, 1985
A;Title: Amino acid sequence of the heavy chain of human alpha-factor XIia (activated Ha
A;Reference number: A22248; MUID:85182674; PMID:3886654
A;Accession: A22248
A;Molecule type: protein
A;Residues: 20-379 <MCN>
R;Fujikawa, K.; McMullen, B.A.
J. Biol. Chem. 258, 10924-10933, 1983
A;Title: Amino acid sequence of human beta-factor XIia.
A;Reference number: A21037; MUID:83291041; PMID:6604055
A;Accession: A21037
A;Molecule type: protein
A;Residues: 354-362, 373-615 <FUJ>
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.
J. Biol. Chem. 267, 5102-5107, 1992
A;Title: O-linked fucose is present in the first epidermal growth factor domain of facto
A;Reference number: A4406; MUID:92184750; PMID:1544894
A;Contents: annotation; carbohydrate binding site
C;Genetics:

A;Gene: GDB:F12
A;Cross-references: GDB:119892; OMIM:234000
A;Map position: 5q34.5qter
A;Introns: 197; 391; 72/2; 96/1; 133/1; 212/1; 212/1; 267/2; 340/1; 417/2; 463/1; 511
C;Complex: factor XII, prekallikrein, and HWW kinogen form a complex bound to anionic
C;Function:
A;Description: factor XIia catalyzes the proteolytic activation of plasminogen, plasma
krein
A;Pathway: blood coagulation; fibrinolysis
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-372,373-615/Product: coagulation factor XIia, alpha form #status experimental <A12
F;47-88/Domain: fibronectin type II repeat homology <FB2>
F;98-130/Domain: EGF homology <EG1>
F;135-170/Domain: fibronectin type I repeat homology <1F1>
F;178-209/Domain: EGF homology <EG2>
F;217-295/Domain: kringle homology <KR>
F;298-356/Region: proline-rich
F;354-362,373-615/Product: coagulation factor XIia, beta form #status experimental <B12
F;373-609/Domain: trypsin homology <TRY>
F;98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-277,266-295
F;109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;299,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F;412,461,563/Active site: His, Asp, Ser #status predicted
-Query Match 32.8%; Score 495; DB 1; Length 615;
Best Local Similarity 38.5%; Pred. No. 4.1e-35;
Matches 107; Conservative 44; Mismatches 109; Indels 18; Gaps 6;
QY 1 KPSPFPPEELK---PQCQ---KTLRPFKIIGGEFTTINQPMFPAIYRRHRGGSVTVC 54
Db 344 KREQPPSLTRNGPLSCGRLKSLSSMTVVVGLVALRGAPHYAALYWGHS-----FC 397
QY 55 GGSIIISCWISATHCFIDPKEDYIVLGRSLNLTQGMKFFVENILHXDYSDT 114
Db 398 AGSLIAFCVWLTAACHLQDRPAEDLTWLVGERRNHSCEPQTFLVRSYRLHEAFS--P 455
QY 115 LAHNDIALKIR-SKEGRCAQPSRTITICLPMSVNDPQFCTSCITGFGKSTNDYLP 173
Db 456 VSYQHDLLALQLQEDADGSSCALLSPYQVCLPSGAARPETTLQCVAGHGHQFEGAREY 515
QY 174 PEQKMTVVKLISHRECQPHYVSEVTTMLCAADPQWKTSCQDGGPLVCSLQ-- 231
Db 516 ASFLQEAQVFPFLSLERCSAPDVHGSSILPGLMCAAGFLEGGTDAQCQDGGPLVCEQAAE 575
QY 232 -RMTLTGIVSWGRCALKDKPGVYTVSHFLPWIRSH 268
Db 576 RRLTLQGIISWVGCGDRNRPVGYTVDAVYLAWIHT 613
RESULT 17
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C;Accession: A46688
R;Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A;Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII.
A;Reference number: A46688; MUID:93252878; PMID:7683665
A;Accession: A46688
A;Molecule type: mRNA
A;Residues: 1-655 <MIY>
A;Cross-references: DDBJ:D14012; NID:G219680; PIDN:BA03113.1; PID:G219681
A;Experimental source: liver (mRNA); serum (protein)
A;Note: sequence extracted from NCBI backbone (NCBI:131227, NCBI:131228)
A;Note: parts of the sequence, including the amino ends of the heavy and light chains,
C;Genetics:
A;Gene: GDB:HGFA; HGFA; HGFAP

A;Cross-references: GDB:9954514
A;Map position: 4p16-4p16
A;Function:
A;Description: activates hepatocyte growth factor by specific proteolytic cleavage
A;Pathway: tissue repair and regeneration
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C;Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F;1-34/Domain: signal sequence #status predicted <SIG>
F;108-148/Domain: fibronectin type II repeat homology <IF2>
F;164-197/Domain: EGF homology <EG1>
F;202-237/Domain: fibronectin type I repeat homology <IF1>
F;245-278/Domain: EGF homology <EG2>
F;286-367/Domain: kringle homology <KRG>
F;373-407/Domain: hepatocyte growth factor activator light chain #status experimental <
F;408-655/Domain: hepatocyte growth factor activator heavy chain #status experimental <
F;408-641/Domain: trypsin homology <TRY>
F;40,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36
F;447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.7%; Score 493.5; DB 1; Length 655;
Best Local Similarity 38.3%; Pred. No. 6e-35;
Matches 105; Conservative 43; Mismatches 109; Indels 17; Gaps 6;

QY 1 KPSPPEELKFOGQK-----TLRPFKIIGGEFTTIENQPFALYYRRHGGSVTYVCG 55
DB 385 EPASPGRQ---ACRRHKKFTLRPR--IIGSSSLPGSHPLAAY---IGDS---FCA 433

QY 56 GSLSISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTL 115
DB 434 GSLVHTCWVVAHCFSHSPRDSVSVLGGHPFNRTDVTQFGIEKIPYTLXSVFNP 493

QY 116 AHNDTALLKIRKEGRCAQPSRTIOTICLPSMYNDPQFGTSCEITGFGKNSDLYLPE 175
DB 494 SDH-DLVRLRLKKGRCATRSFVQICLPBGSTFPAGHKQIAGWHLENVSGYS 552

QY 176 QLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTL 235
DB 553 SLREALVPLVADHKCSPEVYGADISPNMLCAGYFCCKDACQDGGGGLACEKNGVAYL 612

QY 236 TGIIVSWGRGALKDKPGVYTRVSHFLPWIRSHTK 269
DB 613 YGIISWGGCGRLHKPGVYTRVANYVDWINDRIR 646

RESULT 18
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N;Alternate names: Hageman factor
C;Species: Cavia porcellus (guinea pig)
C;Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C;Accession: S28941
R;Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, H.
Biochim. Biophys. Acta 1159, 113-121, 1992
A;Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage si
A;Reference number: S28941; MUID:93003367; PMID:1390917
A;Accession: S28941
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-603 <SEM>
A;Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C;Keywords: hydrolase; serine proteinase
F;46-87/Domain: fibronectin type II repeat homology <IF2>
F;134-169/Domain: fibronectin type I repeat homology <FBI>
F;177-208/Domain: EGF homology <EGF>
F;216-294/Domain: kringle homology <KRG>
F;359-597/Domain: trypsin homology <TRY>

Query Match 32.2%; Score 485.5; DB 2; Length 503;
Best Local Similarity 40.7%; Pred. No. 2.7e-34;
Matches 111; Conservative 41; Mismatches 102; Indels 19; Gaps 8;

QY 6 PBEKFOGQKTIKPRP-----KIIGBETTIENQPFALYYRRHGGSVTYVCGSLISP 61
DB 338 PETSLLCGQR-LRKRLSSLRIVGGVALVLPGAHYIAALY----WGS--NFCGSLIAP 390

QY 62 CWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDI 121
DB 391 CWVLTAAHCLQNRPAPELKVVLQDHRNQCCEHQTILAVHSYRLHEAFSPS--SYLNDL 448

QY 122 ALLKI-RSKEGRCAQPSRTIOTICLPSMYNDPQFG--TSCEITGFGKNSDLYLPEOLK 178
DB 449 ALLRLQKSADGSCAQLSPYVQTVCLPSPGAPPSSETTCCEVAGWGHHQFEGAEYSYFLQ 508

QY 179 MTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTL 235
DB 509 EAQVPLISSEKSCSFEVDGDFLGMLCAGLEGCTDACQDGGPLVCEDEAAEHLIL 568

QY 236 TGIIVSWGRGALKDKPGVYTRVSHFLPWIRSHT 268
DB 569 RGIIVSWGSGCGDRNKPQVYTDVASYLTWIKHT 601

RESULT 19
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N;Alternate names: Hageman factor (activated)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C;Accession: S45281; A61329
R;Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): com
A;Reference number: S45281; MUID:9424782; PMID:8186251
A;Accession: S45281
A;Molecule type: mRNA
A;Residues: 1-593 <SHI>
A;Cross-references: GB:S70164
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as C
is, and ATC for residue 505 as Leu
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).
A;Reference number: A61329; MUID:77182112; PMID:861210
A;Accession: A61329
A;Molecule type: protein
A;Residues: 10-16,'X',18-19;525-550 <FUJ>
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F;37-78/Domain: fibronectin type II repeat homology <IF2>
F;88-120/Domain: EGF homology <EGF>
F;125-160/Domain: fibronectin type I repeat homology <FBI>
F;207-287/Domain: kringle homology <KRG>
F;350-587/Domain: trypsin homology <TRY>
F;541/Active site: Ser #status predicted

Query Match 30.0%; Score 453; DB 2; Length 593;
Best Local Similarity 38.5%; Pred. No. 1.8e-31;
Matches 102; Conservative 44; Mismatches 101; Indels 18; Gaps 7;

QY 13 CGQ---KTLRPFKIIGGEFTTIENQPFALYYRRHGGSVTYVCGSLISPCWISATH 69
DB 336 CGQRLKWLSSLNRVGGVALVLPGAHYIAALYWDQ-----HFCAGSLIAPCWVLTAAH 389

QY 70 CFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIR-S 128
DB 390 CLQNRPAPELKVVLQDHRNQCCEHQTILAVRDLHEAFSPITYQH--DLALVRLQES 447

QY 129 KEGRCAQPSRTIOTICLPSMYNDPQFGTS--CEITGFGKNSDLYLPEOLKMTVVKLIS 186
DB 448 ADGCCAHPSPFPVQVCLPSTTAARPAESEAACEVAGWGHHQFEGGE--YSSPTQEQVPLID 506

QY 187 HRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGLVCSLQGRMTLITGIVSWGR 243

plasma kallikrein (EC 3.4.21.34) precursor - human
 N/Alternate names: kininogenin; plasma prekallikrein
 C/Species: Homo sapiens (man)
 C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 18-Jun-1999
 C/Accession: A00921; A37939
 R/Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
 Biochemistry 25, 2410-2417, 1986
 A/Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four
 A/Reference number: A00921; MUID:86243359; PMID:3521732
 A/Accession: A00921
 A/Molecule type: mRNA
 A/Residues: 1-638 <CHU>
 A/Cross-references: GB:M13143; NID:G190262; PIDN:AAA60153.1; PID:G190263
 R/McMullen, B.A.; Fujikawa, K.; Davie, E.W.
 Biochemistry 30, 2050-2056, 1991
 A/Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of
 A/Reference number: A37939; MUID:91152016; PMID:1998666
 A/Accession: A37939
 A/Molecule type: protein
 A/Residues: 20-27;40-46,'X',48,'H',50,'X',52-70,'X',75-76,'X',78-80;103-113;131-140;141-
 260;283,'X',285;287-291,'X',293-295;314-317,'X',319-320;321-324,'X',329-333;334-339,'X',
 525;538-551;562,'X',564-567;573,'X',575-576;578-583,'X',585;592-604 <MCM>
 C/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex
 C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li-
 are linked by one or more disulfide bonds.
 C/Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal
 ing and may also play a role in the renin-angiotensin system by converting proteinin
 C/Genetics:
 A/Genes: GDB:KLK3
 A/Cross-references: GDB:127575; OMIM:229000
 A/Map position: 4q35-4q35
 C/Superfamily: coagulation factor XI; trypsin homology
 C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-638/Product: plasma kallikrein #status predicted <MAT>
 F:20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>
 F:20-109/Domain: apple repeat <AP1>
 F:110-199/Domain: apple repeat <AP2>
 F:200-289/Domain: apple repeat <AP3>
 F:391-638/Domain: apple repeat <AP4>
 F:391-621/Domain: plasma kallikrein light chain #status predicted <LCH>
 F:391-621/Domain: trypsin homology <TRY>
 F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383
 F:427,308,396,453,494/binding site: carbohydrate (Asn) (covalent) #status experimental
 F:318-347,346-345/Disulfide bonds: status predicted
 F:390-391/Cleavage site: Arg-Ile (coagulation factor XIa) #status predicted
 F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 26.7%; Score 402; DB 1; Length 638;
 Best Local Similarity 33.1%; Pred. No. 5.3e-27;
 Matches 89; Conservative 56; Mismatches 98; Indels 24; Gaps 7;
 QY 17 TURPREKIIIGGEFTTIENQFWFAAIY-----RRHRGGSVTVYVCGSLISPCWVISATHCF 71
 DB 384 TTKTSTRIVGGTNSNGENFWQVSLQVKLTQQRH-----LCGSLGHQWVLTAAHCF 436
 QY 72 IDYPKEDIVVILGRSLNSNGENKPFVENLILKDYSDTLAHNDIALKLRSEK 131
 DB 437 DGLPDDWRIYSGIUNLSDITKDTFSPQIITHQYKVSSEGNH--DIALIKLQAP-- 492
 QY 132 RCAQPSRTITQICLPSPMYNDPQGTSCETITGFG--KENSTDVLYPEQLKMTVVVKLISHRE 189
 DB 493 --LNYTEFQKPLCPKSGKSTIYINCWTGWFSGKEG---IQNILQKNIPLVTHNEE 547
 QY 190 CQOPHYGSEVTVTKMLCAADPOWKTSCGSDGGLVCSLQGRMTLTGIVSGRCAKLD 249
 DB 548 CQK-RQDYTKITQRMVWVCAQYKGGKDGKSDGGLVCKHGMWRLVGTISNGEGCARRE 606
 QY 250 KPGVTVRVSHFLPWRSHTKKEENGLA 275
 DB 607 QGQVTVKRVYWDWLTLEKTKQSDGKA 632

RESULT 26
 S00845
 hepsin (EC 3.4.21.-) - human
 C/Species: Homo sapiens (man)
 C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 18-Jun-1999
 C/Accession: S00845
 R/Leytus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.
 Biochemistry 27, 1067-1074, 1988
 A/Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane di-
 A/Reference number: S00845; MUID:88209431; PMID:2835076
 A/Accession: S00845
 A/Molecule type: mRNA
 A/Residues: 1-417 <LEV>
 A/Cross-references: EMBL:X07732; NID:G32063; PIDN:CAA30558.1; PID:G32064
 C/Genetics:
 A/Genes: GDB:HPN; TMPSRS1; hepsin
 A/Cross-references: GDB:135685; OMIM:142440
 A/Map position: 19q11-19q13.2
 C/Superfamily: hepsin; trypsin homology
 C/Keywords: hydrolase; liver; serine proteinase; transmembrane protein
 F:23-45/Domain: transmembrane #status predicted <TM>
 F:163-400/Domain: trypsin homology <TRY>
 F:188-204,291-359,322-338,349-381/Disulfide bonds: #status predicted
 F:203,257,353/Active site: His, Asp, Ser #status predicted
 Query Match 26.3%; Score 396.5; DB 1; Length 417;
 Best Local Similarity 35.8%; Pred. No. 9.6e-27;
 Matches 101; Conservative 51; Mismatches 89; Indels 41; Gaps 14;
 QY 13 CQKTLRPRFXIIIGGEFTTIENQFWFAAIYRRHRGGSVTVYVCGSLISPCWVISATHCFI 72
 DB 153 CQRRKL-PVDRIVGGRDTSIGRWQVSL--RYDG---AHLCCGSLISGDWVLTAAHCF- 205
 QY 73 DYPKEDIVVILGRSLNSNGE-----MKFEVENLILKDY----SADTLAHNDIAL 123
 DB 206 --PNNR---VLSRVRVFAVAVASPHGLQGVQVYHGGYLPFRDPNSENNDIAL 260
 QY 124 LKIRSEGRCAQPSRTITQICLPSPMYNDPQGTSCETITGFGKENSTDVLYPEQ---LXMT 180
 DB 261 VHLSSP---LPLTEYIOPVCLPAAGQALVDGKICTVTGWS---NTQY-YGQAGVLQEA 312
 QY 181 VKULISHRECOOPHYGSEVTVTKMLCAADPOWKTSCGSDGGLVCSLQGRMTLT 236
 DB 313 RVPILNSDVCNGADFYGNQIKPKFCAGYPEGGIDACQSDGSPFVCDISRTPRMRLC 372
 QY 237 GIVSGRCAKLDKPGVTVRVSHFLPW---IRSHTKKEENGL 274
 DB 373 GIVSWGTCALAAQKPGVTVRVSDFRWFQAIKTHS--EASGM 413

RESULT 27
 S55066
 trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
 N/Alternate names: trypsinogen II
 C/Species: Gallus gallus (chicken)
 C/Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
 C/Accession: S55066; S72347
 R/Wang, K.; Gan, L.; Lee, I.; Hood, L.
 Biochem. J. 307, 471-479, 1995
 A/Title: Isolation and characterization of the chicken trypsinogen gene family.
 A/Reference number: S55065; MUID:95251611; PMID:7733885
 A/Accession: S55066
 A/Molecule type: mRNA
 A/Residues: 1-248 <WAN1>
 A/Cross-references: EMBL:U15157; NID:G603906; PIDN:AAA79914.1; PID:G603907
 A/Experimental source: clone 2-P29
 A/Accession: S72347
 A/Molecule type: DNA
 A/Residues: 1-248 <WAN2>
 A/Cross-references: EMBL:U15157; NID:G603906; PIDN:AAA79914.1; PID:G603907
 A/Experimental source: clone 2-P29
 C/Superfamily: trypsin; trypsin homology
 C/Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen

F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-25/Domain: activation peptide #status predicted <APT>
F:26-246/Product: trypsin II #status predicted <MAT>
F:26-241/Domain: trypsin homology <TRY>
F:65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 26.2%; Score 395.5; DB 2; Length 248;
Best Local Similarity 38.7%; Pred. No. 6.4e-27;
Matches 94; Conservative 40; Mismatches 84; Indels 25; Gaps 7;

QY 23 KIIGGFTTINQWPFAAIYRRHRGSGVTYVCGGSLISPCWVISATHCFIDYPKKEDYIV 82
DB 25 KIVGYTCPEHSYQVSL-----NSGYHFCGSLNSQWVLSAAHCY-----KSRIOV 73

QY 83 YLGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIQT 142
DB 74 RLGEYNIDVEDSEVRSSVRIIRHPYSSITL--NNDIMLIKLAS-----AVEYSADIQ 127

QY 143 ICLPSMYNDPOFGTSCITGFKENSTDYLYPEOLKMTVVKLISHRECOQPHYVGEVTT 202
DB 128 IALPS--SCAKAGTECLISGWNTLSNGYNYPELLOCLNAPILSDQECQEA--YPGDITS 183

QY 203 KMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLP 262
DB 184 NMICVGFLEGGKSCQDSDGPPVNCGE---LQIVSWGIGCALKGYPGVYTKVCNYVD 239

QY 263 WIR 265
DB 240 WIQ 242

RESULT 28
TREBTR
N:Contains: trypsinogen
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change 18-Jul-1997
C:Accession: A90164; A00946; S08774
R:Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
Biochem. Biophys. Res. Commun. 24, 346-352, 1966
A:Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.
A:Reference number: A90164; MUID:67168846; PMID:5967094
A:Accession: A90164
A:Molecule type: Protein
A:Residues: 1-57,'Q',59-67,'Q',69-150,'N',152-176,'N',178-229 <MIK>
R:Hartley, B.S.
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970
A:Reference number: A93755
A:Contents: annotation; revisions
R:Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975
A:Title: Amino acid sequence of dogfish trypsin.
A:Reference number: A00950; MUID:75146445; PMID:1092332
A:Contents: annotation; revisions
A:Note: the sequence agrees with that shown
R:Bode, W.; Schwager, P.
J. Mol. Biol. 98, 693-717, 1975
A:Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution
A:Reference number: A92954; MUID:76072097; PMID:512
A:Contents: annotation; X-ray crystallography; binding sites for calcium, substrate, and
C:Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a termi
s pseudotrypsin. A cleavage may also occur after Arg-105.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-229/Product: trypsinogen #status experimental <ZYM>
F:1-6/Domain: activation peptide #status experimental <APT>
F:7-222/Domain: trypsin homology <TRY>
F:7-131,132-229/Product: alpha-trypsin #status experimental <MPT>
F:6-7/Cleavage site: Lys-1le (enteropeptidase) #status experimental
F:13-143,31-47,115-216,122-168,154-168,179-203/disulfide bonds: #status experimental
F:46,90,183/Active site: His, Asp, Ser #status experimental
F:58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental

F:131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match 26.2%; Score 394.5; DB 1; Length 229;
Best Local Similarity 37.4%; Pred. No. 7.1e-27;
Matches 91; Conservative 41; Mismatches 86; Indels 25; Gaps 8;

QY 23 KIIGGFTTINQWPFAAIYRRHRGSGVTYVCGGSLISPCWVISATHCFIDYPKKEDYIV 82
DB 6 KIVGYTCGANTVPVQVSL-----NSGYHFCGSLNSQWVLSAAHCY-----KSGIOV 54

QY 83 YLGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKRGCAQPSRTIQT 142
DB 55 RLGEDNINWVEGNEQFISAKSVIHPVSYNSNTL--NNDIMLIKLAS-----AASLNSRVAS 108

QY 143 ICLPSMYNDPOFGTSCITGFKENSTDYLYPEOLKMTVVKLISHRECOQPHYVGEVTT 202
DB 109 ISLPT--SCASAGTQCLISGWNTKSTGTSYPDVAKCLKAPILSDSSCKSA--YFGQITS 164

QY 203 KMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLP 262
DB 165 NMFAGYLEGGKDKSCQDSDGPPVCS--GK--LQIVSWGSGCAQKNKPGVYTKVCNYVS 220

QY 263 WIR 265
DB 221 WIK 223

RESULT 29
B61545
N:plasma (EC 3.4.21.7) precursor - sheep (fragments)
A:Alternate names: plasminogen
C:Contains: miniplasminogen
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C:Accession: B61545; S28200
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: B61545
A:Molecule type: Protein
A:Residues: 1-37;38-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A:Title: Complete amino acid sequence of ovine miniplasminogen.
A:Reference number: S28200; MUID:93149995; PMID:1492092
A:Accession: S28200
A:Molecule type: protein
A:Residues: 118-460 <SC2>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase;
F:1-37/Domain: activation peptide (fragment) #status experimental <APT>
F:38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F:41-118/Domain: kringle homology <KR4>
F:118-460/Product: miniplasminogen #status experimental <MIN>
F:132-211/Domain: kringle homology <KRS>
F:226-460/Domain: plasmin chain B #status experimental <BCH>
F:231-453/Domain: trypsin homology <TRY>
F:272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 26.0%; Score 392.5; DB 2; Length 460;
Best Local Similarity 36.0%; Pred. No. 2.4e-26;
Matches 96; Conservative 35; Mismatches 107; Indels 29; Gaps 7;

QY 5 PBEELFQCCQKTLRPR---FKIIGGEFTTINQWPFAAIYRRHRGSGVTYVCGGSLISP 61
DB 209 PQCESSFDGCKPKVEPKKCPARVGGCVATPSHPWQVSLRRSR-----EHFCGGTLISP 264

QY 62 CWVISATHCFIDYPKKEDYIVVLGRSLNSNTQGMKFEVENILHKDYSADTLAHNDI 121
DB 265 EWLTAHCLDSILGSPFYTVILGAHYEMAREASVQEIFVSRFLFLEPSRA-----DI 316

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2004, 12:58:25 ; Search time 18 Seconds

(without alignments)

721.076 Million cell updates/sec

Title: 1508

US-09-880-503-5

Sequence: 1 KPSPPEBLKFCQGGKTLRP.....VSHFLPWIRSHTKBENGLAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1505	99.8	431	1	UROK HUMAN
2	1434.5	95.1	433	1	P00749; Q15844; Q16618; Q969W5; STANDARD; PRT; 431 AA.
3	1255	83.2	442	1	UROK_PIG
4	1188	78.8	433	1	UROK_BOVIN
5	1119.5	74.2	432	1	UROK_RAT
6	1112.5	73.8	434	1	UROK_MOUSE
7	688.5	45.7	434	1	UROK_CHICK
8	583	38.7	562	1	TPA HUMAN
9	570.5	37.8	559	1	TPA RAT
10	565	37.5	566	1	TPA_BOVIN
11	564.5	37.4	559	1	TPA_MOUSE
12	547	36.3	431	1	URTE_DESRO
13	546	36.2	394	1	URTG_DESRO
14	546	36.2	477	1	URT2_DESRO
15	543	36.0	477	1	URT1_DESRO
16	497.5	33.0	653	1	HGFA_MOUSE
17	495	32.8	615	1	FA12_HUMAN
18	493.5	32.7	655	1	HGFA_HUMAN
19	485.5	32.2	603	1	FA12_CAVPO
20	453	30.0	593	1	FA12_BOVIN
21	442.5	29.3	875	1	NETR_HUMAN
22	425.5	28.2	761	1	NETR_MOUSE
23	423.5	28.1	418	1	HATT_HUMAN
24	408.5	27.1	343	1	PSS8_HUMAN
25	407	27.0	638	1	KAL_MOUSE
26	405	26.9	277	1	KLKD_HUMAN
27	402.5	26.7	436	1	HEPS_MOUSE
28	402	26.7	638	1	KAL_HUMAN
29	398.5	26.4	455	1	TMS5_MOUSE
30	396.5	26.3	417	1	HEPS_HUMAN
31	395.5	26.2	248	1	TRY3_CHICK
32	394.5	26.2	243	1	TRY1_BOVIN
33	392.5	26.0	343	1	PLMN_SHEEP

34	390.5	25.9	247	1	TRY2_BOVIN
35	390	25.9	263	1	CTR2_CANFA
36	389.5	25.8	244	1	KLK6_HUMAN
37	388.5	25.8	453	1	TMS3_MOUSE
38	388	25.7	457	1	TMS5_HUMAN
39	386.5	25.6	342	1	PSS8_RAT
40	386.5	25.6	416	1	HEPS_RAT
41	385	25.5	638	1	KAL_RAT
42	384.5	25.5	311	1	TRYG_MOUSE
43	383.5	25.4	790	1	PLMN_PIG
44	382	25.3	263	1	CTR8_RAT
45	380.5	25.2	342	1	PSS8_MOUSE

ALIGNMENTS

RESULT :

UROK_HUMAN
ID UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W5;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator)
GN PLA0.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heynaker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiranatsu R., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=85203359; PubMed=3888571;
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Boilen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feigold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Villalón D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A., RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., RA Scherch A., Schein J.E., Jones S.J.M., Marra W.A., RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). RN [7] SEQUENCE OF 66-431 FROM N.A. RP MEDLINE=8427206; PubMed=6589620; RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Biasi F.; RT "Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA." RT Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731 (1984). RN [8] SEQUENCE OF 21-177. RP MEDLINE=83055084; PubMed=6754569; RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E., RA Flohe L.; RT "The primary structure of high molecular mass urokinase from human urine. The complete amino acid sequence of the A chain." RT Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165 (1982). RN [9] SEQUENCE OF 156-176 AND 179-224. RP MEDLINE=83003608; PubMed=6749491; RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W., RA Studer R.O.; RT "Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence data of the two polypeptide chains." RT Eur. J. Biochem. 125:251-257 (1982). RN [10] SEQUENCE OF 158-410. RP MEDLINE=83055099; PubMed=6754572; RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.; RT "The complete amino acid sequence of low molecular mass urokinase from human urine." RT Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058 (1982). RN [11] X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS). RP MEDLINE=96000858; PubMed=8591045; RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D., RA Dobson C.M., Stuart D.I., Jones E.V.; RT "The crystal structure of the catalytic domain of human urokinase-type plasminogen activator." RT Structure 3:681-691 (1995). RN [12] X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411. RP MEDLINE=20266327; PubMed=10805774; RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G., RA Bode W., Magdolen V., Huber R., Moroder L.; RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly selective inhibitors of human urokinase." RT Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118 (2000). RN [13] STRUCTURE BY NMR. RP MEDLINE=89127526; PubMed=2536903; RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.; RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR." RT Nature 337:579-582 (1989). RN [14] STRUCTURE BY NMR OF 67-155. RP MEDLINE=93003110; PubMed=1327118; RA Li X., Smith R.A.G., Dobson C.M.;

"Sequential 1H NMR assignments and secondary structure of the kringle domain from urokinase." Biochemistry 31:9562-9571 (1992). RN [15] STRUCTURE BY NMR OF 67-155. RP MEDLINE=94149701; PubMed=8107091; RA Li X., Bokman A.M., Linaas M., Smith R.A.G., Dobson C.M.; RT "Solution structure of the kringle domain from urokinase-type plasminogen activator." RT J. Mol. Biol. 235:1548-1559 (1994). RN [16] VARIANT LEU-141. RP MEDLINE=96186279; PubMed=8652631; RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K., RA Sawasaki Y., Hanada K.; RT "Characterization of single chain urokinase-type plasminogen activator with a novel amino-acid substitution in the kringle structure." RT Biochim. Biophys. Acta 1293:83-89 (1996). RN [17] VARIANT LEU-141. RP MEDLINE=97218551; PubMed=9065988; RA Conne B., Berczy M., Belin D.; RT "Detection of polymorphisms in the human urokinase-type plasminogen activator gene." RT Thromb. Haemost. 77:434-435 (1997). RN [18] ERRATUM. RA Conne B., Berczy M., Belin D.; RT Thromb. Haemost. 78:973-973 (1997). RN [19] VARIANT LEU-141. RP MEDLINE=97337920; PubMed=9194591; RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Heil W., RA Creutzburg S., Graeff H., Magdolen V.; RT "Mutational analysis of the genes encoding urokinase-type plasminogen activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer." RT Electrophoresis 18:686-689 (1997). CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR THERAPY OF THROMBOLYTIC DISORDERS. CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin. CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN. CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis. CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. CC -!- SIMILARITY: Contains 1 kringle domain. CC -!- SIMILARITY: Contains 1 EGF-like domain. CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch). CC EMBL; X02419; CAA26268.1; - DR EMBL; M15476; AAA61253.1; - DR EMBL; D00244; BAA00175.1; - DR EMBL; D11143; BAA01919.1; - DR EMBL; X02760; CAA26535.1; - DR EMBL; AF377330; AAK53822.1; - DR EMBL; BC013575; AAH13575.1; - DR EMBL; K03266; AAC97138.1; - DR EMBL; K02286; AAA61252.1; - DR EMBL; A21571; CAA01559.1; - DR EMBL; A18397; CAA01390.1; - DR PIR; A00931; UKHU. DR PDB; 1KDU; 31-OCT-93.

Query Match 99.8%; Score 1505; DB 1; Length 431;
 Best Local Similarity 99.8%; Pred. No. 9.6e-129;
 Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPEELKFCQCKTLPRFKIIGGETTIENQPFAAIYRRHRGSGSVTVYCGSLIS 60
 DB 156 KPSSPEELKFCQCKTLPRFKIIGGETTIENQPFAAIYRRHRGSGSVTVYCGSLIS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFKENSTDYLYPEQLKMT 335

QY 181 VKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSGGGLVCSLQGRMTLTGIVS 240
 DB 336 VKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSGGGLVCSLQGRMTLTGIVS 395

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 431

RESULT 2
 UROK_PAPCY STANDARD; PRT; 433 AA.

ID UROK_PAPCY STANDARD; PRT; 433 AA.
 AC P16237;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
 DE (U-plasminogen activator).
 GN PLAU
 OS Papio cynocephalus (Yellow baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheinae; Papio.
 OX NCBI_TaxID=9556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thoracic aorta;
 RX MEDLINE=90287734; PubMed=2113276;
 RA Au Y.P.T., Wang T.W., Clowes A.W.;
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
 RT plasminogen activator."
 RL Nucleic Acids Res. 18:3411-3411(1990).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-val bond in
 CC plasminogen to form Plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by, and for commercial
 entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).

 CC EMBL; X51935; CAA36200.1; -
 CC PIR; S14687; UKBAY.
 CC HSSP; P00749; 1LWV.
 CC MEROPS; S01.231; -
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF_like.

DR InterPro; IPR006210; IBSF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; TYP_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
 FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
 FT DOMAIN 26 62 EGF-LIKE.
 FT DOMAIN 69 150 KRINGLE.
 FT DOMAIN 151 177 CONNECTING PEPTIDE.
 FT DOMAIN 178 433 SERINE PROTEASE.
 FT DISULFID 30 38 BY SIMILARITY.
 FT DISULFID 32 50 BY SIMILARITY.
 FT DISULFID 52 61 BY SIMILARITY.
 FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 208 224 BY SIMILARITY.
 FT DISULFID 216 287 BY SIMILARITY.
 FT DISULFID 315 384 BY SIMILARITY.
 FT DISULFID 347 363 BY SIMILARITY.
 FT DISULFID 374 402 BY SIMILARITY.
 FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
 FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
 FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
 FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
 SQ SEQUENCE 433 AA; 48595 MW; 816D22DFDDC8792 CRC64;

Query Match 95.1%; Score 1434.5; DB 1; Length 433;
 Best Local Similarity 92.8%; Pred. No. 2.3e-122;
 Matches 259; Conservative 13; Mismatches 4; Indels 3; Gaps 1;

QY 1 KPSSPEELKFCQCKTLPRFKIIGGETTIENQPFAAIYRRHRGSGSVTVYCGSLIS 60
 DB 155 KPSSPEELKFCQCKTLPRFKIIGGETTIENQPFAAIYRRHRGSGSVTVYCGSLIS 214

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 215 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 274

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQ---FGTSCITGFKENSTDYLYPEQL 177
 DB 275 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPNDPPFGTSCITGFKENSTDYLYPEQL 334

QY 178 KMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSGGGLVCSLQGRMTLTG 237
 DB 335 KMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSGGGLVCSLQGRMTLTG 394

QY 238 IVSWGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
 DB 395 IVSWGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 433

RESULT 3
 UROK_PIG
 ID UROK_PIG
 AC P04185;

STANDARD; PRT; 442 AA.

FT	DISULFID	220	235	BY SIMILARITY.
FT	DISULFID	228	299	BY SIMILARITY.
FT	DISULFID	324	393	BY SIMILARITY.
FT	DISULFID	356	372	BY SIMILARITY.
FT	DISULFID	383	411	BY SIMILARITY.
FT	ACT_SITE	235	235	CHARGE RELAY SYSTEM.
FT	ACT_SITE	286	286	CHARGE RELAY SYSTEM.
FT	ACT_SITE	387	387	CHARGE RELAY SYSTEM.
FT	ACT_SITE	241	241	Q -> H (IN REF. 1; CAA25806).
FT	CONFLICT	242	242	Q -> H (IN REF. 1; CAA26511).
FT	CONFLICT	288	288	A -> GS (IN REF. 1; CAA25806).
FT	CONFLICT	442	442	AA; 49116 MW; EE32FCEE501321EE CRC64;
FT	SEQUENCE	442	442	AA; 49116 MW; EE32FCEE501321EE CRC64;
QY	Query Match	83.2%	Score 1255;	DB 1; Length 442;
QY	Best Local Similarity	82.1%	Pred. No. 4.2e-106;	
QY	Matches	225;	Conservative 23;	Mismatches 26; Indels 0; Gaps 0;
Db	168	PFS	PEKVEFCQCGKALRPRFKIIGEGFTTIENQWFAAIYRHRGGSVTVYCGSLISP	22
QY	62	CWV	SATHCFIDYPKKEDYIVYLGSRLSNSNTQGEKMPVENLIHKKOYSADTLAHHNDI	121
Db	228	CWV	SATHCFINYOQKEDYIVYLGROTLSSHTGEMKFEVKLILHEDYSADSLAHHNDI	287
QY	122	ALLKIRSKGRCAQPSRTI	QICLPISMYNDPFGTSCETIGFGKENSTDYLYPEOLKMTV	181
Db	298	ALLKIRTDKGCQACPSS	QICLPVNGDAHFASCEIVGFGKEDPSDYLYPEOLKMTV	347
QY	182	VKLASHRECQPHYGSEV	TTKMLCAADPOWKTDSCQSGGGPLVCSLOGRMTLTGIVSW	241
Db	348	VKLASHRECQPHYGSEV	TTKMLCAADPOWKTDSCQSGGGPLVCSLOGRMTLTGIVSW	407
QY	242	GGCALDKDPVYTRVSH	FLPWRSHTKKEENGIA	275
Db	408	GGCALDKDPVYTRVSR	FLTWIHTHVGGENGLA	441
RESULT 4	UROK_BOVIN	STANDARD;	PRT;	433 AA.
ID	UROK_BOVIN	Q05589;	Q28209;	
AC	01-FEB-1994	(Rel. 28, Created)		
DT	01-FEB-1994	(Rel. 28, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Urokinase-type plasminogen activator precursor	(EC 3.4.21.73)	(uPA)	
DE	(U-plasminogen activator).			
GN	PLAU			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
NCBI	TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Arctic endothelium;			
RC	MEDLINE=9316119; PubMed=8385052;			
RA	Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,			
RA	Schleuning W.-D.;			
RT	"Bovine urokinase-type plasminogen activator and its receptor:			
RT	cloning and induction by retinoic acid.";			
RL	Gene 125:177-183(1993).			
RN	[2]			
RP	SEQUENCE OF 12-433 FROM N.A.			
RC	TISSUE=Kidney;			
RA	Ravn P., Berglund L., Petersen T.B.;			
RT	"Cloning and characterization of the bovine plasminogen activators uPA			
RT	and tPA.";			
RL	Int. Dairy J. 5:605-617(1995).			
CC	-I- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in			
CC	plasminogen to form plasmin.			
CC	-I- INDUCTION: By retinoic acid.			
CC	-I- SIMILARITY: BELONGS TO PERTIDASE FAMILY S1.			


```

DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02400; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (AI) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 382 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 74.2%; Score 1119.5; DB 1; Length 432;
Best Local Similarity 71.4%; Pred. No. 7.6e-94;
Matches 197; Conservative 39; Mismatches 39; Indels 1; Gaps 1;

QY 1 KSSPPPEELKFCGOKTLPREKLIIGETTITENQWFAAIYRRHGGG-VTVVCGGSLI 59
Db 156 KPSSTVDQGGFCGQKALPRFKIVGGERTVVENQWFAAIYLRKNGGSPPFKCGGSLI 215
QY 60 SPQWVIASTHCTIDYPKKEDYIVYLGSRNLNGNTQGMKFVEVENLILKDYSDATLAHNN 119
Db 216 SPQWVASATHCFVQPKKEEYVYVYLGSKNSYNPGEMKFVEVEQLILHEDFSDETLAFEN 275
QY 120 DTALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFQTSCEITGPKENSTLYLPEQLKM 179
Db 276 DTALLKIRSTGCAQPSRTIQTICLPPRFQDAPFGSDCEITGPKESATDYFPKDLKM 335
QY 180 TVVKLIHRECOQPHYVYGGSEVTKMLCAADPQWKTDSCGDSGGPLVCSLOGRMTLTGIV 239
Db 336 SVVKLIHQKQKPHYVYGGSEINVKMLCAADPQWKTDSCGDSGGPLICNDGRPTLSGIV 395
QY 240 SMGRGALKDKPGVYTRVSHFLPWIRSHYKENGLA 275
Db 396 SMGSGCAEKNGVYTRVSHFLPWIRSHYKENGLA 431

RESULT 6
UROK_MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN FLAU.
OS Mus musculus (Mouse).

```

```

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85179474; PubMed=2985383;
RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88163499; PubMed=2831940;
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279(1987).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02389; CAA26231.1; -
DR EMBL; M17922; AAA40539.1; -
DR PIR; A29420; UKMS.
DR HSSP; P00749; 1XDU.
DR MEROPS; S01-231; -.
DR MGD; MGI:97611; Plau.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. Protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02400; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (AI).
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.

```


RESULT 8
 TPA_HUMAN
 ID TPA_HUMAN STANDARD; PRT; 562 AA.
 AC P00750; Q15103;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).
 GN PLAT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Melanoma;
 RX MEDLINE=83115262; PubMed=6337343;
 RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
 RA Ward C.A., Bennett W.E., Yelverton E., Seeburg P.H., Heyneker H.L.,
 RA Goeddel D.V., Collen D.;
 RT "Cloning and expression of human tissue-type plasminogen activator
 RT cDNA in E. coli.";
 RL Nature 301:214-221(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal lung;
 RX MEDLINE=89828259; PubMed=3133640;
 RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
 RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
 RT from human fetal lung cells.";
 RL Nucleic Acids Res. 16:5695-5695(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88054470; PubMed=2824147;
 RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
 RA Hsiung N.;
 RT "Expression of human uterine tissue-type plasminogen activator in
 RT mouse cells using BPV vectors.";
 RL DNA 6:461-472(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86196143; PubMed=3009482;
 RA Frierer Degen S.J., Rajput B., Reich E.;
 RT "The human tissue plasminogen activator gene.";
 RL J. Biol. Chem. 261:6972-6985(1986).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84298137; PubMed=6089198;
 RA Ny T., Elgh F., Lund B.;
 RT "The structure of the human tissue-type plasminogen activator gene:
 RT correlation of intron and exon structures to functional and
 RT structural domains.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86284200; PubMed=3030401;
 RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
 RA Opdenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
 RT "Cloning of cDNA coding for human tissue-type plasminogen activator
 RT and its expression in Escherichia coli.";
 RL Mol. Biol. Med. 3:279-292(1986).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC TISSUE=Umbilical vein;
 RX MEDLINE=90192129; PubMed=2107528;
 RA Siebert P.D., Fong K.;
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
 RT human endothelial cells.";
 RL Nucleic Acids Res. 18:1086-1086(1990).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Watley J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP SEQUENCE OF 212-361 FROM N.A.
 RX MEDLINE=83169656; PubMed=6572897;
 RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
 RA Josephson S.;
 RT "Isolation of cDNA sequences coding for a part of human tissue
 RT plasminogen activator.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
 RN [10]
 RP SEQUENCE OF 1-36 FROM N.A.
 RX MEDLINE=85289338; PubMed=3161893;
 RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
 RA Schleuning W.-D.;
 RT "Isolation and characterization of the human tissue-type plasminogen
 RT activator structural gene including its 5' flanking region.";
 RL J. Biol. Chem. 260:11223-11230(1985).
 RN [11]
 RP SEQUENCE OF 31-562 FROM N.A.
 RX MEDLINE=91291340; PubMed=1368681;
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuoka S., Higashio K.;
 RT "Purification and characterization of tissue plasminogen activator
 RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
 RL Agric. Biol. Chem. 55:1225-1232(1991).
 RN [12]
 RP SEQUENCE OF 36-562.
 RC TISSUE=Melanoma;
 RX MEDLINE=8500468; PubMed=6433976;
 RA Pohl G., Kaelstroom M., Bergsdorf N., Wallen P., Joernvall H.;
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
 RT derived amino acid sequence, identify the active site serine residue,
 RT establish glycosylation sites, and localize variant differences.";
 RL Biochemistry 23:3701-3707(1984).
 RN [13]
 RP SEQUENCE OF 33-52 AND 311-330.
 RC TISSUE=Melanoma;
 RX MEDLINE=83209620; PubMed=6682760;
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
 RT "Purification and characterization of a melanoma cell plasminogen
 RT activator.";
 RL Eur. J. Biochem. 132:681-686(1983).
 RN [14]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=90092112; PubMed=2513186;
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
 RT "Carbohydrate structure of recombinant human uterine tissue
 RT plasminogen activator expressed in mouse epithelial cells.";
 RL Eur. J. Biochem. 186:273-286(1989).
 RN [15]
 RP CARBOHYDRATE-LINKAGE SITE THR-96.
 RX MEDLINE=91159408; PubMed=1900431;
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;

QY	191	QQPHYYSSEVTTXMLCAAD-----PQWKT-DSQQDSGGPLVCSLQGRMTLTGIIVSWGRG	244
DB	477	TSCHLLNRTVTDNMLCAGTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG	536
QY	245	CALKDKPGVYTRVSHFLPWIRSHTK	269
DB	537	CGQKQVPGYTKVNTYLDWIRDNR	561

RESULT 9			
TPA_RAT	ID	TPA_RAT	STANDARD; PRT; 559 AA.
AC	AD	FI9637;	
DT	DT	01-FEB-1991 (Rel. 17, Created)	
DT	DT	01-NOV-1991 (Rel. 20, Last sequence update)	
DT	DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (t-PA)	
DE	DE	(t-PA) (t-plasminogen activator).	
GN	GN	PLAT.	
OS	OS	Rattus norvegicus (Rat).	
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OC	OC	NCBI_TaxId=10116;	
RN	RN	[1]	
RP	RP	SEQUENCE FROM N.A.	
RX	RX	MEDLINE=89170114; PubMed=3148445;	
RX	RX	Medline=89170114; PubMed=3148445;	
RT	RT	NY T.; Leonardson G.; Heuen A.J.W.;	
RT	RT	"Cloning and characterization of a cDNA for rat tissue-type	
RT	RT	plasminogen activator.";	
RL	RL	DNA 7:671-677(1988).	
RL	RL	[2]	
RP	RP	SEQUENCE FROM N.A.	
RX	RX	MEDLINE=90130448; PubMed=2105315;	
RX	RX	Feng P.; Ohlsson M.; Ny T.;	
RT	RT	"The structure of the TATA-less rat tissue-type plasminogen activator	
RT	RT	gene. Species-specific sequence divergences in the promoter predict	
RT	RT	differences in regulation of gene expression.";	
RL	RL	J. Biol. Chem. 265:2022-2027(1990).	
CC	CC	-!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMAGEN PLASMINOGEN	
CC	CC	TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY	
CC	CC	CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT	
CC	CC	ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND	
CC	CC	MANY OTHER PHYSIOPATHOLOGICAL EVENTS.	
CC	CC	-!- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in	
CC	CC	plasminogen to form plasmin.	
CC	CC	-!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE	
CC	CC	BOND.	
CC	CC	-!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.	
CC	CC	-!- FTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER	
CC	CC	PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER	
CC	CC	ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.	
CC	CC	-!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A	
CC	CC	CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.	
CC	CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	CC	-!- SIMILARITY: Contains 1 EGF-like domain.	
CC	CC	-!- SIMILARITY: Contains 1 fibronectin type I domain.	
CC	CC	-!- SIMILARITY: Contains 2 kringle domains.	
CC	CC	-----	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	CC	use by non-profit institutions as long as its content is in no way	
CC	CC	modified and this statement is not removed. Usage by and for commercial	
CC	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	CC	or send an email to license@isb-sib.ch).	
CC	CC	-----	
CC	CC	EMBL; M23697; AAA41812.1; -	
DR	DR	EMBL; M31187; AAA42261.1; -	
DR	DR	EMBL; M31185; AAA42261.1; JOINED.	
DR	DR	EMBL; M31186; AAA42261.1; JOINED.	
DR	DR	EMBL; M31187; AAA42261.1; JOINED.	
DR	DR	EMBL; M31188; AAA42261.1; JOINED.	

	Query Match	38.7%	Score 583;	DB 1;	Length 562;
	Best Local Similarity	44.9%	Pred. No. 3.9e-45;		
	Matches 119;	Conservative 39;	Mismatches 97;	Indels 10;	Gaps 5;
QY	13	CG-OXLTLPREKFIIGTGFTTIENQPFPAALYRSH-RGGSVTVVCGGSLISPCWVISATHC	70		
Db	299	CLGRQYSGQFRIKGGFLADIAHPWQALFAXHRRSPGERFLCGGILISSCWILSAHC	358		
QY	71	FIDVPKEDYIVYLGSRLSNNTQGENMKFEVENLIHKDYSADTLAHNDIALKIRSE	130		
Db	359	FQERAPPHLTVILGRYTVVPVEEQKFVEKYIVHKEFDODT-YNDIALQLKSDS	416		
QY	131	GRCAQPERTIOTICLPSMYNDPQFGSCETGTGKENSIDYLPQLKMTVVKLISHREC	190		
Db	417	SRCAQESVVRTCLPPADQLQDPDWTCELSUGKHEALSPYSERLKEARVPYSRC	476		

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC HSSP; P00750; LRTF. 1.
 CC MEROPS; S01.232.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR000083; Fibronectin.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00039; fni; 1.
 CC Pfam; PF00051; Kringle; 2.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000395; Kringle; 2.
 CC SMART; SM00181; EGF; 1.
 CC SMART; SM00058; FNI; 1.
 CC SMART; SM00130; KR; 2.
 CC SMART; SM00020; TRYPSIN; 1.
 CC PROSITE; PS00022; EGF 1; 1.
 CC PROSITE; PS01186; EGF 2; 1.
 CC PROSITE; PS01253; FIBRONECTIN 1; 1.
 CC PROSITE; PS00021; KRINGLE 1; 1.
 CC PROSITE; PS00070; KRINGLE 2; 2.
 CC PROSITE; PS02400; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Plasma; Kringle; EGF-like domain; Repeat; Signal.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 FT SIGNAL 1 21
 FT PROPEP 22 33
 FT CHAIN 34 566
 FT CHAIN 34 314
 FT CHAIN 315 566
 FT CHAIN 40 82
 FT DOMAIN 83 121
 FT DOMAIN 128 209
 FT DOMAIN 219 300
 FT DOMAIN 315 566
 FT ACT_SITE 361 361
 FT ACT_SITE 410 410
 FT ACT_SITE 517 517
 FT DISULFID 42 72
 FT DISULFID 70 79
 FT DISULFID 87 98
 FT DISULFID 92 109
 FT DISULFID 111 120
 FT DISULFID 128 209
 FT DISULFID 149 191
 FT DISULFID 180 204
 FT DISULFID 219 300
 FT DISULFID 240 282
 FT DISULFID 271 295
 FT DISULFID 303 434
 FT DISULFID 346 362
 FT DISULFID 346 423
 FT DISULFID 448 523
 FT DISULFID 480 496
 FT DISULFID 513 541
 FT CARBOHYD 153 153
 FT CARBOHYD 487 487
 SQ SEQUENCE 566 AA; 63701 MW; 2E86BB8432276C3 CRC64;
 Query Match 37.5%; Score 565; DB 1; Length 566;
 Best Local Similarity 43.4%; Pred. No. 1.7e-43;

Matches 115; Conservative 41; Mismatches 99; Indels 10; Gaps 4;
 QY 13 CG-QKTLRPRFKIIGGFTTIENQWFAAIY-RRHGGSVTVYCGSLSPCWISATHC 70
 DB 303 CGLRQYKRPQPRIKGLFADITSPWQAAIFVQVRRSPGERFLCGGILISSCWLSAAHC 362
 QY 71 FIDYPKXEDYIVLGRSLNSNTQGENKFEVENILHKDYSADTLAHNDIALLKIRSK 130
 DB 363 FOERYPPHLKVLGRYRLVPGEEQTFEVEKIIHKFPDDDT--YNDIALHLKSDS 420
 QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKNSDYLYPEQLKMTYVKLIHREC 190
 DB 421 LTCARESASVRTICLPDASLQLPDWTCELSGYCHSSSPFSERLKEAHLVLPSSRC 480
 QY 191 QOPHYGSEVTTKMLCAADPQW-----KTDSCGDSGGPLVCSLQGRMTLTGIVSGRG 244
 DB 481 TSQHLFNRTVTNNMLCAGDTRSGGDHTNLHDACQDSGGPLVCMKDNHMTLVGIISWGLG 540
 QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
 DB 541 CGRKDVPGVYTKVTNYLDWDINDTR 565
 RESULT 11
 ID TPA_MOUSE STANDARD; PRT; 559 AA.
 AC P11214; O91VP2;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)
 DE (t-PA) (t-plasminogen activator).
 GN PLAT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88087303; PubMed=2826484;
 RA Rickles R.J., Darrow A.L., Strickland S.;
 RT "Molecular cloning of complementary DNA to mouse tissue plasminogen
 RT activator mRNA and its expression during f9 teratocarcinoma cell
 RT differentiation.";
 RL J. Biol. Chem. 263:1563-1569(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=223825; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOGEN PLASMINOGEN
 CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 22
 ID AAR66253 standard; protein; 393 AA.
 AC AAR66253;
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M20.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX Synthetic.

OS Key Location/Qualifiers
 FH Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..372
 FT /label= X1
 FT Region 373..393
 FT /label= Y1

PN DE4323754-Cl.
 XX 01-DEC-1994.
 XX 15-JUL-1993; 93DE-4323754.
 XX 15-JUL-1993; 93DE-4323754.
 PR (CHEF) GRUENENTHAL GMBH.
 PA Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX Example 1; Page 10 and Fig 1; 34pp; German.

CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs, which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 393 AA;

Query Match 100.0%; Score 1508; DB 16; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.le-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKXLRPRFKIIGGEFTTIENQWPFAAIYRHRGSGVTVYCGSLIS 60
 DB 90 KPSSPPEELKFCQCKXLRPRFKIIGGEFTTIENQWPFAAIYRHRGSGVTVYCGSLIS 149

QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENILHKDYSADTLAHND 120
 DB 150 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENILHKDYSADTLAHND 209

QY 121 IALKIRSKGRCACQPSRTTIQTICLPSMYNDPQGTSCETITGFGKENSIDYLYPEQLKMT 180
 DB 210 IALKIRSKGRCACQPSRTTIQTICLPSMYNDPQGTSCETITGFGKENSIDYLYPEQLKMT 269

QY 181 VKLISHRECOQPHYGYSEVTTWMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 240
 DB 270 VKLISHRECOQPHYGYSEVTTWMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 23
 AAR99885
 ID AAR99885 standard; peptide; 393 AA.
 AC AAR99885;
 XX 27-JAN-1997 (first entry)
 DT M36: fibrinolytic and anticoagulant activity contg. protein.
 DE Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen;
 KW urokinase; activator; streptokinase; staphylokinase; APSAC;
 KW anisolated plasminogen streptokinase activator complex; hirudin;
 KW hirullin; antistatin; pWrt27; pWS1; pSE8; pW56.
 XX Synthetic.
 OS EP712934-A2.
 PN 22-MAY-1996.
 XX 03-NOV-1995; 95EP-0117316.
 XX 17-NOV-1994; 94DE-4440892.
 PR (CHEF) GRUENENTHAL GMBH.
 PA Heinzl-Wieland R, Steffens GJ, Wnendt S;
 PI WPI; 1996-240720/25.
 XX Proteins with fibrinolytic and anticoagulant activity - useful as
 PT thrombolytic agents
 XX Disclosure; Fig 18; 59pp; German.

CC New peptides (I) with fibrinolytic and anticoagulant activity
 CC comprise a plasminogen-activating amino acid sequence (A) fused
 CC at the N- and/or C-terminus to a thrombin and/or factor Xa
 CC inhibiting amino acid sequence (B). Excluded from the claims
 CC are (i) where (A) is Ser47 to Leu411 of unglycosylated urokinase
 CC linked at the C-terminus to sequences (i) to (iii):
 CC T1-RP-T2-CGGNGDGFEBIPEYL-T3 (i)
 CC T1-RP-LLNPNDKYEPFWEDEKNE (ii)
 CC T1-RSSEFEFEIDEEEK (iii)
 CC Where T1= P or V; T2= L or a bond; T3= Q or OH.
 CC (A) is pref. (pro)urokinase; tissue plasminogen activator (tPA),
 CC bat-PA (all opt. modified by deletion, substitution, insertion and/or
 CC addn.); streptokinase; staphylokinase; and/or APSAC (anisolated

CC plasminogen streptokinase activator complex), esp. prourokinase
 CC (411 amino acids) or its Ser17 or Ser138 to Leu411
 CC fragments, or t-PA (527 amino acids) or its Ser9Arg to 527Pro or
 CC 174Ser to 527Pro fragments.
 CC (B) has hirudin or hirudin activities; or is derived from the human
 CC thrombin receptor, antistatin and/or the tick anticoagulant peptide.
 CC Most pref. are the 65 amino acid hirudin sequence or one of the six
 CC sequences given in AAR99879 to AAR99884.
 CC Plasminids pWJ27 (M51), pW51 (M5112), pSE8 (M36) and pW56 (M43)
 CC contain the sequences encoding AAR99885 to AAR99888, respectively.
 CC The products were tested in human citrated plasma (5 microg in 200
 CC microl 1:10 diluted plasma). The thrombin time was then 1.2, 3, 2.8
 CC and 1.2 times greater, respectively, than in the absence of the product.
 XX Sequence 393 AA;
 SQ

Query Match 100.0%; Score 1508; DB 17; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGGSLIS 60
 DB 91 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGGSLIS 150
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
 DB 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 210
 QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180
 DB 211 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 270
 QY 181 VVKLISHRECQPPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 271 VVKLISHRECQPPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 330
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 331 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 366

RESULT 24
 AAR99596
 ID AAR99596 standard; Protein; 393 AA.
 XX
 AC AAR99596;
 XX
 DT 05-DEC-1996 (first entry)
 XX
 DE Chimeric protein M37 encoded by pSE9.
 XX
 KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 KW plasminogen activating sequence; fibrinolysis; infarction;
 KW angina pectoris; deep vein thrombosis.
 XX
 OS Synthetic.
 XX
 PN EP714982-A2.
 XX
 FD 05-JUN-1996.
 XX
 PF 16-NOV-1995; 95EP-0118050.
 XX
 PR 30-NOV-1994; 94DE-4442665.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Janocha E, Steffens GJ, Wnendt S;
 XX
 DR WPI; 1996-269715/28.
 XX
 PT Chimeric protein contg. plasminogen activating sequence and
 PT thrombin-inhibiting sequence - useful as thrombus-specific

PT thrombolytic agent with rapid action
 XX
 XX Example 1; Page 19-20; 37pp; German.
 CC
 CC Example 1 describes the prodn. of plasmids pSE1 and pSE9
 CC contg. a DNA encoding a chimeric protein with fibrinolytic and
 CC thrombin-inhibiting properties.
 CC pSE1 encodes the protein given in AAR99597 and pSE9 encodes
 CC the protein given in AAR99596.
 XX
 XX Sequence 393 AA;
 SQ

Query Match 100.0%; Score 1508; DB 17; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGGSLIS 60
 DB 91 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGGSLIS 150
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
 DB 151 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 210
 QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180
 DB 211 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 270
 QY 181 VVKLISHRECQPPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 271 VVKLISHRECQPPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 330
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 331 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 366

RESULT 25
 AAR99597
 ID AAR99597 standard; Protein; 393 AA.
 XX
 AC AAR99597;
 XX
 DT 05-DEC-1996 (first entry)
 XX
 DE Chimeric protein M38 encoded by pSE1.
 XX
 KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
 KW plasminogen activating sequence; fibrinolysis; infarction;
 KW angina pectoris; deep vein thrombosis.
 XX
 OS Synthetic.
 XX
 PN EP714982-A2.
 XX
 PD 05-JUN-1996.
 XX
 PF 16-NOV-1995; 95EP-0118050.
 XX
 PR 30-NOV-1994; 94DE-4442665.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Janocha E, Steffens GJ, Wnendt S;
 XX
 DR WPI; 1996-269715/28.
 XX
 PT Chimeric protein contg. plasminogen activating sequence and
 PT thrombin-inhibiting sequence - useful as thrombus-specific
 PT thrombolytic agent with rapid action
 XX
 XX Example 1; Page 21-22; 37pp; German.
 XX

CC Example 1 describes the prodn. of plasmids pSE1 and pSE9
 CC contg. a DNA encoding a chimeric protein with fibrinolytic and
 CC thrombin-inhibiting properties.
 CC pSE1 encodes the protein given in AAR99597 and pSE9 encodes
 CC the protein given in AAR99596.
 XX
 SQ Sequence 393 AA;

Query Match 100.0%; Score 1508; DB 17; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60
 DB 91 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 150
 QY 61 PCWISATHCFIDYPKEDYVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 120
 DB 151 PCWISATHCFIDYPKEDYVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 210
 QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 180
 DB 211 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 270
 QY 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 240
 DB 271 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 330
 QY 241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 331 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 366

RESULT 26
 AAR66262
 ID AAR66262 standard; protein; 395 AA.

AC AAR66262;
 XX
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 XX Bifunctional urokinase variant M29.
 DE
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 KW
 XX Synthetic.

Key Location/Qualifiers
 FT Region 1..365 /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..371
 FT Region /label= X1
 FT 372..395
 FT Region /label= Y1

PN DE4323754-C1.

PD 01-DEC-1994.

XX 15-JUL-1993; 93DE-4323754.

XX

PR 15-JUL-1993; 93DE-4323754.
 XX (CHEF) GRUENTHAL GMBH.
 PA Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX
 XX WPI; 1995-015191/03.
 XX
 PT New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 395 AA;

Query Match 100.0%; Score 1508; DB 16; Length 395;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60
 DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 149
 QY 61 PCWISATHCFIDYPKEDYVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWISATHCFIDYPKEDYVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 180
 DB 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 269
 QY 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 27
 AAR66265
 ID AAR66265 standard; protein; 395 AA.

AC AAR66265;
 XX
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 XX Bifunctional urokinase variant M32.
 DE
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Region 1..365 /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85

QY 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSVDLYPEQLKMT 180
 DB 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSVDLYPEQLKMT 269
 QY 181 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 29
 AAR66248
 ID AAR66248 standard; protein; 397 AA.
 AC AAR66248;
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 XX Bifunctional urokinase variant M15.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..378
 FT Region 379..397
 FT /label= X1
 FT /label= Y1
 XX DE4323754-C1.
 XX
 XX 01-DEC-1994.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Heinkel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wendt S;
 PI
 DR WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derives and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 10 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula

CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 397 AA;
 Query Match 100.0%; Score 1508; DB 16; Length 397;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPEELKFCQCKTLPRFKIIGGETTIENQFWFAAIYRRHGGSVYVCGSLIS 60
 DB 90 KPSSPEELKFCQCKTLPRFKIIGGETTIENQFWFAAIYRRHGGSVYVCGSLIS 149
 QY 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSVDLYPEQLKMT 180
 DB 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSVDLYPEQLKMT 269
 QY 181 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 30
 AAP50871
 ID AAP50871 standard; protein; 411 AA.
 XX
 XX AAP50871;
 DT 30-NOV-1991 (first entry)
 DE Sequence encoded by cDNA sequence for human urokinase zymogen
 DE (Japanese Patent Application No.37119/84).
 XX Thrombolytic agent; plasminogen activator activity; fibrin affinity;
 KW enzyme.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site 158..159
 FT /note= "potential cleavage site which generates
 FT the two-chain form from the zymogen"
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX
 XX EP139447-A.
 XX
 XX 02-MAY-1985.
 XX
 XX 07-SEP-1984; 84EP-0306117.
 XX
 XX 17-OCT-1983; 83JP-0195051.

PR 13-SEP-1983; 83JP-0170354.
 XX (GREG) GREEN CROSS CORP.
 PA Tsukada M, Tanaka K, Iga Y;
 XX WPI; 1989-101389/14.
 XX Fibrinolytic activity enhancer -
 XX comprising plasminogen for enhancing activity of single-chain
 PT pro-urokinase without causing systemic fibrinolysis
 XX Disclosure; page 7; 8pp; English.

XX Zymogen AAP50871 is the inactive precursor form of human urokinase.
 CC Urokinase zymogen is cleaved into the two-chain form composed of
 CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular
 CC wt. of 20,000) chains when treated with catalytic amounts of plasmin.
 CC The patenters claim a new urokinase zymogen which has mol. wt. ca.
 CC 50,000, a single chain molecular structure, and selective affinity
 CC for fibrin. It is a thrombolytic agent which manifests its
 CC plasminogen activator activity on cleavage by proteolytic enzymes
 CC (e.g. plasmin) and has higher affinity for fibrin than known forms
 CC of urokinase.

XX Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 6; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVCGSLIS 60
 Db 136 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVCGSLIS 195
 QY 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 196 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMT 180
 Db 256 IALLKIRSGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMT 315
 QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 316 VVKLISHRECCQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 276
 Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 411

RESULT 31
 AAP93589
 ID AAP93589 standard; protein; 411 AA.
 XX AAP93589;
 XX 19-JUN-1990 (first entry)
 XX Amino acid sequence of single-chain pro-urokinase extracted from human
 DE renal cells.
 DE Single-chain prourokinase; fibrinolytic activity enhancer;
 XX plasminogen; thrombosis.
 KW Homo sapiens.
 XX EP310065-A.
 XX 05-APR-1989.
 XX 29-SEP-1988; 88EP-0116067.
 XX 01-OCT-1987; 87JP-0248937..

XX (GREG) GREEN CROSS CORP.
 XX Tsukada M, Tanaka K, Iga Y;
 XX WPI; 1989-101389/14.
 XX Fibrinolytic activity enhancer -
 XX comprising plasminogen for enhancing activity of single-chain
 PT pro-urokinase without causing systemic fibrinolysis
 XX Disclosure; page 7; 8pp; English.
 XX It was extracted from human renal cells and purified using a monoclonal
 CC antibody for single-chain prourokinase (scPU) recovered from hybridoma
 CC cells created from the fusion of mouse myeloma cells with mouse BAUS/c
 CC spleen cells which had been preliminarily immunised with scPU. The
 CC purified product has a molecular weight of 54,000 in SDS/polyacrylamide
 CC gel electrophoresis. The combined use of scPU and plasminogen at a
 CC specific ratio can enhance fibrinolytic activity of scPU without causing
 CC any systemic fibrinolysis. This enables lowering the dose of scPU which
 CC might relieve side effects. They can be used for the treatment of
 CC thrombosis and obstructive diseases.

SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 10; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVCGSLIS 60
 Db 136 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVCGSLIS 195
 QY 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 196 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMT 180
 Db 256 IALLKIRSGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMT 315
 QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 316 VVKLISHRECCQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 276
 Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 411

RESULT 32
 AAP96146
 ID AAP96146 standard; protein; 411 AA.
 XX AAP96146;
 XX 03-OCT-2002 (updated)
 DT 21-JAN-1991 (first entry)
 XX Sequence encoded by entire prourokinase (PKU) gene from PKU-producing
 DE tumour cell line ATCC CCL138 clone pUC20.
 XX Thrombosis; fibrinolytic agent; venous disease;
 KW arterial disease therapy.
 XX Unidentified.
 XX EP312941-A.
 XX 26-APR-1989.
 XX 15-OCT-1988; 88EP-0117186.

```
XX PR 23-OCT-1987; 87DE-3735917.
XX PA (BADI ) BASF AG.
XX FI Koerwer W, Kurfurst M, Baldinger V, Doerper T, Schwarz M;
XX DR WPI; 1989-123847/17.
XX DR N-PSDB; AAN91617.
XX PT New N-shortened pro-urokinase peptide cpds. with thrombolytic activity -
XX PT and longer in vivo half life, opt. with replacement of
XX PT arginine-156
XX PS Example; Fig 2, p 705-8; 21pp; German.
XX CC New N-shortened pro-urokinase peptide cpds. were prepd. from pUC20. The
XX CC new peptides are useful for treating venous and arterial occlusive
XX CC diseases.
XX CC (Updated on 03-OCT-2002 to add missing OS field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 1508; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.2e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVVCGSLIS 60
DB 136 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVVCGSLIS 195
QY 61 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 196 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
QY 121 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
DB 256 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 315
QY 181 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
RESULT 33
AAP91684
XX ID AAP91684 standard; Protein; 411 AA.
XX AC AAP91684;
XX XX Location/Qualifiers
FH Disulfide-bond 19..11
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 102..126
FT Disulfide-bond 71..113
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
XX OS Homo sapiens.
XX FH Key
XX FT Disulfide-bond 19..11
XX FT Disulfide-bond 13..31
XX FT Disulfide-bond 33..42
XX FT Disulfide-bond 50..131
XX FT Disulfide-bond 102..126
XX FT Disulfide-bond 71..113
XX FT Disulfide-bond 148..279
XX FT Disulfide-bond 189..205
XX KW Plasminogen activator; urokinase; thrombosis therapy;
XX KW cardiovascular disorder; myocardial infarct.
XX OS Homo sapiens.
```

```
FT Disulfide-bond 197..268
FT FT Disulfide-bond 293..361
FT FT Disulfide-bond 325..341
FT FT Disulfide-bond 352..380
FT FT Modified-site 302
FT FT /label= glycosylation site
XX XX
XX WO8911531-A.
XX PN 30-NOV-1989.
XX PD
XX XX
XX PF 04-MAY-1989; 89WO-US01947.
XX XX
XX PR 20-MAY-1988; 88US-0196909.
XX XX
XX PA (GETH ) GENENTECH INC.
XX XX
XX PI Anderson S, Keyt B;
XX XX
XX DR WPI; 1989-370725/50.
XX XX
XX PT New plasminogen activator variants with additional glycosylation -
XX PT having increased circulating half life in plasma, used for treating
XX PT cardiovascular disorders
XX PS Disclosure; Fig 3; 60pp; English.
XX CC The new plasminogen activator (PA) variants of the invention are
XX CC derived from human urokinase, pro-urokinase or esp. tissue PA (tPA).
XX CC They are useful for treating cardiovascular disorders such as
XX CC myocardial infarct or thrombosis. Typical doses are 0.3 mg/kg for
XX CC infarct and 0.1-0.2 mg/kg for thrombosis.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 1508; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.2e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVVCGSLIS 60
DB 136 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVVCGSLIS 195
QY 61 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 196 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
QY 121 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
DB 256 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 315
QY 181 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 316 VVKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
RESULT 34
AAP06244
XX ID AAP06244 standard; Protein; 411 AA.
XX XX
XX AC AAP06244;
XX XX
XX DT 07-DEC-1990 (first entry)
XX XX
XX DE Urokinase precursor protein.
XX XX
XX KW Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
XX KW myocardial infarction;
```

XX OS Homo sapiens.
 XX PN EP380334-A.
 XX PD 01-AUG-1990.
 XX PF 25-JAN-1990; 90EP-0300772.
 XX PR 17-MAY-1989; 89JP-0121405.
 XX PR 27-JAN-1989; 89CP-0016406.
 XX PA (GREC) GREEN CROSS CORP.
 XX PI Matsuda H, Ueda Y, Tananouchi K;
 XX DR WPI; 1990-233117/31.
 XX PT Urokinase precursor-lipid composite - used as thrombolytic agent,
 PT having prolonged half-life in the blood, enhanced
 PT bioavailability and improved activity
 XX Claim 3; Fig 1; 11pp; English.
 XX By forming a precursor-lipid composite, the half-life of this
 CC thrombolytic agent in the blood may be increased, exhibiting
 CC improved activity without abnormal acceleration of fibrinolytic
 CC activity. Compound is useful as a thrombolytic agent in
 CC treatment of cerebral thrombosis, myocardial infarction etc.
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 1508; DB 11; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCKQTLRPRFKIIGGEFTTIENQPFWFAAIYRHRGGSVTVCGGSLIS 60
 DB 136 KPSSPPEELKFCQCKQTLRPRFKIIGGEFTTIENQPFWFAAIYRHRGGSVTVCGGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSKGRCQAQPSRTIQTICLPMSYNDPFGTSCEITGFGKENSTDYLYPEQLKMT 180
 DB 256 IALLKIRSKGRCQAQPSRTIQTICLPMSYNDPFGTSCEITGFGKENSTDYLYPEQLKMT 315
 QY 181 VVKLISHRECQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 316 VVKLISHRECQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
 DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
 RESULT 35
 AAR07902
 ID AAR07902 standard; protein; 411 AA.
 XX AAR07902;
 XX AC
 XX DT 21-FEB-1991 (first entry)
 XX DE Human pro-urokinase variant.
 XX KW Thrombin; fibrin; bleeding; pHR22.
 XX OS Homo sapiens.
 XX PH Key Location/Qualifiers
 FT Domain 10..42

FT Active-site /label= Epidermal growth factor (EGF) domain
 FT 22..24 /label= Modified site
 FT Region 10..19 /label= First loop
 FT Region 20..31 /label= Second loop
 FT Region 33..42 /label= Third loop
 XX EP398362-A.
 XX 22-NOV-1990.
 XX 18-MAY-1990; 90EP-0109473.
 XX 18-MAY-1989; 89JP-0126434.
 XX (GREC) GREEN CROSS CORP.
 XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M;
 XX Kawabe H, Arimura H;
 XX WPI; 1990-350147/47.
 XX N-PSDB; AAQ06133.
 XX Human pro-urokinase variant - produced by recombinant methods,
 PT showing increased half life in blood and high affinity for fibrin.
 XX Disclosure; Fig 1; 27pp; English.
 XX Modified pro-urokinase has a longer half-life in blood, and dissolves
 CC thrombin without causing the spontaneous bleeding associated with
 CC urokinase. The modification puts an epidermal growth factor domain
 CC into the protein, containing the sequence Asn-X-Ser or Asn-X-Thr
 CC where X is any residue.
 CC Plasmid pHR22 is disclosed as containing the modified sequence.
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 1508; DB 11; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCKQTLRPRFKIIGGEFTTIENQPFWFAAIYRHRGGSVTVCGGSLIS 60
 DB 136 KPSSPPEELKFCQCKQTLRPRFKIIGGEFTTIENQPFWFAAIYRHRGGSVTVCGGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSKGRCQAQPSRTIQTICLPMSYNDPFGTSCEITGFGKENSTDYLYPEQLKMT 180
 DB 256 IALLKIRSKGRCQAQPSRTIQTICLPMSYNDPFGTSCEITGFGKENSTDYLYPEQLKMT 315
 QY 181 VVKLISHRECQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 316 VVKLISHRECQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
 DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
 RESULT 36
 AAR07903
 ID AAR07903 standard; protein; 411 AA.
 XX AAR07903;
 XX AC
 XX DT 21-FEB-1991 (first entry)
 XX

RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63989; AAA31594.1; -
CC PIR; JS0599; JS0599.
CC HSP; P98119; IAS1.
CC MEROPS; S01.239; -
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF00089; Trypsin; 1.
CC PRINTS; PRO0722; CHYMOTRYPSIN.
CC PRINTS; PRO0018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KS; 1.
CC SMART; SM00020; TRYD_SPC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPsin_DOM; 1.
CC PROSITE; PS00134; TRYPsin_HIS; 1.
CC PROSITE; PS00135; TRYPsin_SER; 1.
CC Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431
FT DOMAIN 37 75
FT DOMAIN 82 163
FT DOMAIN 179 431
FT ACT_SITE 226 226
FT ACT_SITE 275 275
FT ACT_SITE 362 362
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
FT DISULFID 103 145
FT DISULFID 134 158
FT DISULFID 168 229
FT DISULFID 211 227
FT DISULFID 219 288
FT DISULFID 313 388
FT DISULFID 345 361
FT DISULFID 378 406
FT CARBOHYD 139 139
FT CARBOHYD 352 352

SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;
Query Match 36.3%; Score 547; DB 1; Length 431;
Best Local Similarity 42.9%; Pred. No. 5.1e-42;
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;
QY 4 SPPELKFQCG-QKTLRPRFKIIGGETTIENTNQWFAAIYRRHGGG-VTVVCGSLISP 61
DB 159 SVPVCSKATCGRLRYKPEQLHSTGCLFTDITSHPWAAIFAQNRSSGERELFCGGLISS 218
QY 62 CWVISATHCFID-YPKXEDYIVLGRSLNSNTQEMKFEVENILHKDYSADTLAHND 120
DB 219 CWLTAARHCFQERYPPQHLRV-LGRIVYKPGKEQTFEVEKCIIEEPDDT--YND 275
QY 121 IALLKIRSKGRCAQPSRTQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLMT 180
DB 276 IALLQLKSGSPQCAQESDRAICLPEANLQLPDWTCELSGYGKHSSSPFYSEQLKEG 335
QY 181 VVKLISHRECQPHYGVSEVTKVLCADPQWKT-----DSCQDGGPLVCSLQGRMT 234
DB 336 HVRLYPSRCTSKFLFNKTVNNMLCAGDTRSGEIVYVNDACQDGGPLVCVNDNMT 395
QY 235 LTGIVSWGRCALKXDPGVYTRVSHFLPWIRSHTK 269
DB 396 LLGIISWGVGCGEXDIRGVYTKVTNVLGWIRDNMR 430
RESULT 13
URTG DESRO
ID URTG DESRO STANDARD; PRT; 394 AA.
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSBA
DE gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_Taxid=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RA MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SMO0181; EGF; 1.
DR SMART; SMO0058; FN1; 1.
DR SMART; SMO0130; KR; 1.
DR SMART; SMO0020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477
FT DOMAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 225 477
FT DOMAIN 272 272
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 73 BY SIMILARITY.
FT DISULFID 77 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 214 345 BY SIMILARITY.
FT DISULFID 257 273 BY SIMILARITY.
FT DISULFID 359 434 BY SIMILARITY.
FT DISULFID 391 407 BY SIMILARITY.
FT DISULFID 424 452 BY SIMILARITY.
FT CARBOHYD 185 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 36.2%; Score 546; DB 1; Length 477;
Best Local Similarity 42.9%; Pred. No. 7; le-42;
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;

QY 4 SPPEELKFOG-QKTLRPRFKIGGFETTINQNPWFAIYRHRGGS-VTVVCGGSLISP 61
Db 205 SVPVCSKATGLRKYKEPQLHSGG:FTDITSHPMQAAIFAQNRSSGGRFLCGILISS 264
QY 62 CWVISATHCFID-YPKEDYIVVLGRSLNSNTQGMKEVENLILKDYSDTLAHND 120
Db 265 CWVLTAAHCFQERYPQHRLRV-LGRTYRVKPKKEQTEVEVKXIVHEFDDET--YNN 321
QY 121 IALLKIRKEGRCAQPSRTIQTICLPMSYNDPFGTSCBITGFKENSDYLYPEQLKMT 180
Db 322 IALLQLKSGSPQCAQESDSVRAICLPEANLQLPDWTECLSGYGHKSSSPYSQLEK 381
QY 181 VKLISHRCQDPHYGSEVTTKMLCAADPQWT-----DSCQDSGGLVCSIQGRMT 234
Db 382 HVRLYPSSRTSKFLFNKFTVNNMLCAGTSGEYIPNVHDAQCQDSGGLVCMNDNMT 441
QY 235 LTGIVSWGRCALKDKPGVYTVSVHFLPWIRHTK 269
Db 442 LUGIISWVGCGEKDIPGYTKVTWYLGWIRNMR 476

RESULT 15
URTI DESRO
ID URTI DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSFA
alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Salivary gland;
RC MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
CHARACTERIZATION.
RP MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
Schleuning W.D., Bode W.;
RT "Catalytic domain structure of vampire bat plasminogen activator: a
molecular paradigm for proteolysis without activation cleavage.";
RL Biochemistry 36:13483-13493(1997).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
THE PRESENCE OF FIBRIN I.
CC -!- SUBUNIT: Monomer.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; M63987; AAA31591.1; --
CC EMBL; M63986; AAA31592.1; --
CC PIR; JS0597; JS0597.
CC PDB; 1ASI; 23-MAR-99.
CC MEROPS; S01.232; --
CC GlycoSuiteDB; P98119; --
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF-like.

DR InterPro; IPR000083; Fibrinectn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; fn1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE.
FT DOMAIN 225 477 SERINE PROTEASE.
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM.
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM.
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 214 345 BY SIMILARITY.
FT DISULFID 257 273 BY SIMILARITY.
FT DISULFID 265 334 BY SIMILARITY.
FT DISULFID 359 434 BY SIMILARITY.
FT DISULFID 391 407 BY SIMILARITY.
FT DISULFID 424 452 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC...).
FT CARBOHYD 398 398 N-LINKED (GLCNAC...).
FT TURN 214 215 /Ftid-CAR_000027.
FT STRAND 223 224 N-LINKED (GLCNAC...).
FT TURN 226 227 /Ftid-CAR_000028.
FT STRAND 230 231
FT HELIX 234 236
FT TURN 238 239
FT STRAND 240 245
FT STRAND 254 263
FT TURN 264 265
FT STRAND 266 269
FT HELIX 271 273
FT TURN 280 282
FT STRAND 284 287
FT TURN 297 298
FT STRAND 300 309
FT TURN 311 312
FT TURN 315 317
FT TURN 319 320
FT STRAND 323 328
FT STRAND 338 338
FT TURN 339 340

FT STRAND 341 341
FT STRAND 345 345
FT TURN 349 350
FT TURN 355 356
FT STRAND 358 363
FT STRAND 366 366
FT STRAND 374 374
FT STRAND 379 385
FT HELIX 388 390
FT TURN 393 398
FT TURN 403 404
FT STRAND 405 409
FT TURN 425 426
FT TURN 428 429
FT STRAND 431 436
FT TURN 437 438
FT STRAND 439 448
FT TURN 455 456
FT STRAND 459 463
FT HELIX 464 467
FT HELIX 468 474
SQ SEQUENCE 477 AA; 53616 MW; AA06FD1739C10E5E CRC64;
Query Match 36.0%; Score 543; DB 1; Length 477;
Best Local Similarity 42.2%; Pred.No.1.3e-41;
Matches 117; Conservative 42; Mismatches 98; Indels 20; Gaps 6;
QY 4 SPBELKFCQC-QKTLRPRPKIIGGEFTTIENQFWFAAIYRRHRGGS-VTVVCGGSLISP 61
DB 205 SVPVCSKATCGLRKYKBPQLHSTGGLFTDITSHWQAIPAQRNRSSGERFLCGGILISS 264
QY 62 CWVISATHCFIDYPKKEDI-----VYLGRSLNSNTQGMKFEVENILILHKYSAOTLA 116
DB 265 CWLTAARCF-----QESYLPDQLKVLGRTRYRVKPGEEQTFKVKYIVHKEFDDDT-- 317
QY 117 HNDIALLKIRSKGRCAQPSRTIOTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQ 176
DB 318 YNNDIALLQLKSDSPQCAQSDSVRAICLPANQLPDWTECELSGKGKXSSFFSEQ 377
QY 177 LKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKT-----DSCQDGGGPLVCSLQ 230
DB 378 LKEGHVRLYPSSRCAPKFLFNKTVNNMLCAGDTRSGEIEYPNVHDACQDGGGGLVCMND 437
QY 231 GRMLTIGVSWGRGALKKDKGVVTRYSHFLPWIRSH 267
DB 438 NHMTLLGIISWVGCGEKDVGVTYKVTNYLGIWRDN 474
RESULT 16
HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9UKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGFA)
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1] SEQUENCE FROM N.A.
RN [2] STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2] SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Ruan Y.;
RT "Activation of HGF by endogenous HGF activator is required for

RT metanephric kidney morphogenesis in vitro.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
 CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
 CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
 CC DISULFIDE BOND (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
 CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF099017; AF02489.1; -
 CC EMBL; AF224724; AAF34712.1; -
 CC HSSP; P00763; IDEO.
 CC MEROPS; S01.228; -
 CC MGD; MGI:1859281; Hgfac.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR006209; EGF_Like.
 CC InterPro; IPR000093; Fibronctn1.
 CC InterPro; IPR000562; FN_Type_II.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00008; EGF_2.
 CC Pfam; PF00039; fn1; 1.
 CC Pfam; PF00040; fn2; 1.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00013; ENTPEPIL.
 CC PRINTS; PR00018; KRINGLE.
 CC ProDom; PD000995; FN_Type_II; 1.
 CC ProDom; PD000395; Kringle; 1.
 CC SMART; SM00181; EGF; 2.
 CC SMART; SM00059; FN2; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00022; EGF_1; 2.
 CC PROSITE; PS01186; EGF_2; 1.
 CC PROSITE; PS01253; FIBRONECTIN_1; 1.
 CC PROSITE; PS00023; FIBRONECTIN_2; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS00070; KRINGLE_2; 1.
 CC PROSITE; PS02040; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 29 BY SIMILARITY.
 FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
 FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
 FT CHAIN
 FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
 FT CHAIN
 FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
 FT DOMAIN 157 195 EGF-Like 1.
 FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
 FT DOMAIN 238 276 EGF-Like 2.
 FT DOMAIN 283 364 KRINGLE
 FT DOMAIN 406 653 SERINE PROTEASE.

FT ACT_SITE 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 105 BY SIMILARITY.
 FT DISULFID 119 BY SIMILARITY.
 FT DISULFID 145 BY SIMILARITY.
 FT DISULFID 161 BY SIMILARITY.
 FT DISULFID 166 BY SIMILARITY.
 FT DISULFID 183 BY SIMILARITY.
 FT DISULFID 185 BY SIMILARITY.
 FT DISULFID 194 BY SIMILARITY.
 FT DISULFID 199 BY SIMILARITY.
 FT DISULFID 225 BY SIMILARITY.
 FT DISULFID 234 BY SIMILARITY.
 FT DISULFID 242 BY SIMILARITY.
 FT DISULFID 247 BY SIMILARITY.
 FT DISULFID 266 BY SIMILARITY.
 FT DISULFID 267 BY SIMILARITY.
 FT DISULFID 283 BY SIMILARITY.
 FT DISULFID 304 BY SIMILARITY.
 FT DISULFID 335 BY SIMILARITY.
 FT DISULFID 392 BY SIMILARITY.
 FT DISULFID 430 BY SIMILARITY.
 FT DISULFID 438 BY SIMILARITY.
 FT DISULFID 533 BY SIMILARITY.
 FT DISULFID 565 BY SIMILARITY.
 FT DISULFID 581 BY SIMILARITY.
 FT DISULFID 592 BY SIMILARITY.
 FT CARBOHYD 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 G -> W (IN REF. 2).
 SQ SEQUENCE 653 AA; 70567 MW; 8884B2025DF7FDC CRC64;
 Query Match 33.0%; Score 497.5; DB 1; Length 653;
 Best Local Similarity 38.5%; Pred. No. 2.5e-37;
 Matches 105; Conservative 44; Mismatches 109; Indels 15; Gaps 6;
 QY 2 PSSPPEELKFOGQK-----TLRPRFKIIIGGEFTIENQPFALYYRHRGSGVTYVCGG 56
 Db 382 PESAP-AVRPTCGKRHKRKTFLRPR--IIIGSSSLPGSHPLAAIY---IGNS---FCAG 432
 QY 57 SLISPCWVISATHGCFIDYPKKEDYIVLGRSLRSLNTGEMKFEVENILHKKYDSADTLA 116
 Db 433 SLVHTCWVSAHCFANSPRDSITVLGQHFFRTDVTDTGIEKYVPVTLVSFNP 492
 QY 117 HNDIALKIRSKGRCAQPSRTTQICLPSMYNDPFGTSCEITGFGKENSTLYLPEQ 176
 Db 493 NH-DLVLTIRLKKGERCAVRSQFVQICLPEAGSFPFGTHKQCQIAGWGHMDENVSSYNS 551
 QY 177 LKMTVVKLISHRECOQPHYGVSEVTTKMLCRADPQWKTDSCQDGGGGLVCSLOGRWTLT 236
 Db 552 LLEALVPLVADHKCSSPEVYGDISPMLCAGYFDCSDACQDGGGGLVCKNGVAVLY 611
 QY 237 GIVSWGRCALKDKPGVYTVRVSHFELPWIRSHTK 269
 Db 612 GIISWGDGGRLNKPGVYTVRVANVVDINDIR 644
 RESULT 17
 ID FA12_HUMAN
 AC P00748; P78339;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cosignation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAP).
 GN FN12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN NCBi_taxid=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88007593; PubMed=2888762;

RA Cool D.E., McGillivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region.";
RL J. Biol. Chem. 262:13662-13673(1987).
RN [2].
RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3].
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi R., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL Nucleic Acids Res. 14:3146-3146(1986).
RN [4].
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGillivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa.";
RL J. Biol. Chem. 260:13666-13676(1985).
RN [5].
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RL Biochemistry 25:1525-1528(1986).
RN [6].
RP SEQUENCE OF 20-379
RX MEDLINE=85182674; PubMed=3885654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
RN [7].
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=4604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa.";
RL J. Biol. Chem. 258:10924-10933(1983).
RN [8].
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=36133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material
RT negative patients.";
RL Hum. Mol. Genet. 4:1235-1237(1995).
RN [9].
RP CARBOHYDRATE-LINKAGE SITE THR-109.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [10].
RP VARIANT WASHINGTON D.C. SER-590.
RX MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571->Ser substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN [11].
RP VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;

RA Hovenga J.K., Schaller J., Stricker H., Wullemijn W.A., Furlan M.,
RA Lammle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353->Pro leading to loss of a
RT kallikrein cleavage site.";
RL Blood 84:1173-1181(1994).
RN [12].
RP VARIANT TENRI CYS-53.
RX MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
RA "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT deficiency, occurs through a proteasome-mediated degradation.";
RL Blood 93:4300-4308(1999).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- PTM: O- AND N-GLYCOSYLATED.
CC -!- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR
CC XIIa ACTIVATES FACTOR XI TO FACTOR Xla.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31315; AAB70225.1; -.
DR EMBL; AF538691; AAB67932.1; -.
DR EMBL; M11723; AAA51986.1; -.
DR EMBL; M17466; AAB59490.1; -.
DR EMBL; M17464; AAB59490.1; JOINED.
DR EMBL; M17465; AAB59490.1; JOINED.
DR EMBL; M13147; AAB70224.1; -.
DR EMBL; U71274; AAB51203.1; -.
DR PIR; A29411; KFHU12.
DR HSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR Genew; HGNC:3530; F12.
DR MIM; 234000; -.
DR GO; GO:0003805; F: blood coagulation factor XI activity; TAS.
DR GO; GO:0003806; F: blood coagulation factor XII activity; TAS.
DR GO; GO:0007596; P: blood coagulation; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR006210; IGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.

DR SMART; SMC00181; EGF; 2.
DR SMART; SMC0058; FN1; 1.
DR SMART; SMC0059; FN2; 1.
DR SMART; SMC0030; KR; 1.
DR SMART; SMC0020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW Polymorphism; Disease mutation.
FT SIGNAL 1 19
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.
FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
FT DOMAIN 94 131 EGF-LIKE 1.
FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
FT DOMAIN 174 210 EGF-LIKE 2.
FT DOMAIN 217 295 KRINGLE.
FT DOMAIN 296 349 PRO-RICH.
FT DOMAIN 373 615 SERINE PROTEASE.
FT CARBOHYD 109 109 O-LINKED (FUC).
FT CARBOHYD 249 249 N-LINKED (GLNAC.).
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match 32.8%; Score 495; DB 1; Length 615;
Best Local Similarity 38.5%; Pred. No. 4e-37;
Matches 107; Conservative 44; Mismatches 109; Indels 18; Gaps 6;

QY 1 KPSSPEELK---FOCGO---KTLRFRFKIIGFEFTIENQPFALYRRHGGSVTVYC 54
Db 344 KREQPSLTNGPLSCGQRKLSNTRVVGGLVALRGHPYIAALYWGHS-----FC 397

QY 55 GGSLSIPCWVISATHCFIDYFKKEDIYVILGRSLNSNTQGMKPEVENLIHLKDYSDT 114
Db 398 AGSLIAPCWLTAAHCLQDRPAEDLTVLGQERRNHSCEPCOTLAVRSYRLHEAFS--P 455

QY 115 LAHNDIALKIR-SKEGRCAQPSRIQICLPSMTNDFQGTSCBITFGKENSTDYLY 173
Db 456 VSYCHDLALLRLQEDADGSCALLSPYVQVPLFSGAARFSETTLCQVAGWGHPGAEAY 515

QY 174 PEOLKMTVVKLIHRECOQPHYVGSVTTKMLCAADPQWKTSCQDGGPLVCSLQG-- 231
Db 516 ASFLQAPQVFFLSLRCSAPDVHGSSILPCLMCAQFLGCTGACQDSCGPLVCDQAAE 575

QY 232 -RWLTGIVSWGKCALCKPKGVYTVSHPLPWIRSH 268
Db 576 RRLTQGIISWGGCGDRNKPQVYTDVAYLAWIREHT 613

RESULT 18
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF

DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028 (1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D14012; BAA03113.1; -
EMBL; Z69923; CAA93803.1; -
PIR; A46688; A46688.
HSSP; P00763; 1DPO.
MEROPS; S01.228; -
Genew; HGNC:4894; HGFA.
MIM; 604582; -
GO; GO:0005576; C:extracellular; TAS.
GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR000742; EGF_2.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibrinctn.
InterPro; IPR000562; FN_Type_II.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00008; EGF; 2.
Pfam; PF00039; fn1; 1.
Pfam; PF00040; fn2; 1.
Pfam; PF00051; kringle; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; FNTYPEII.
PRINTS; PR00018; KRINGLE.
ProDom; PD000995; FN_Type_II; 1.
SMART; SM00181; EGF; 2.
SMART; SMC0058; FN1; 1.
SMART; SMC0059; FN2; 1.

```
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372
FT CHAIN 373 407
FT CHAIN 408 655
FT CHAIN 108 148
FT DOMAIN 160 198
FT DOMAIN 200 240
FT DOMAIN 241 279
FT DOMAIN 286 367
FT DOMAIN 408 655
FT ACT_SITE 447 447
FT ACT_SITE 497 497
FT ACT_SITE 598 598
FT DISULFID 108 133
FT DISULFID 122 148
FT DISULFID 164 175
FT DISULFID 169 186
FT DISULFID 188 197
FT DISULFID 202 230
FT DISULFID 228 237
FT DISULFID 245 256
FT DISULFID 250 267
FT DISULFID 269 278
FT DISULFID 286 367
FT DISULFID 307 349
FT DISULFID 338 362
FT DISULFID 394 521
FT DISULFID 432 448
FT DISULFID 440 510
FT DISULFID 535 604
FT DISULFID 567 583
FT DISULFID 594 622
FT CARBOHYD 48 48
FT CARBOHYD 290 290
FT CARBOHYD 468 468
FT CARBOHYD 492 492
FT CARBOHYD 546 546
FT CONFLICT 644 644
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;

Query Match 32.78; Score 493.5; DB 1; Length 655;
Best Local Similarity 38.38; Pred. No. 5.9e-37;
Matches 105; Conservative 43; Mismatches 109; Indels 17; Gaps 6;

Qy 1 KPSSPPPELKFQCGQK-----TLRPREKIIIGBEFTTIENQFWFAAIYRRHSGSVTVVCG 55
Db 385 EPASPGRQ---ACGRHKXRTFLRPR--IIGSSSLPGSHPLAAIY---IGDS---FCA 433
Qy 56 GSLISPCWVISAHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYGSADTL 115
Db 434 GSLVHTCWVVAACHFSHPRSVSVVLGQHFNRTDVTQFGIEKYIPTYLYSVFNP 493
Qy 116 AHNHDIALLKIRSEKGRCAQPSRTIQTICLPMSYNDPQFGTSCEITGFGKNGSTDVLYPE 175
Db 494 SDH-DLVLRILKXKGDRCATESQFVQICLPFGSIFPAGHKCQIAGWGHLDENVSGYSS 552
Qy 176 QLKMVTVKLISHREKQPHYGEVTTKMLCAADPQWKVTSQCGDGGPLVCSLQGRMTL 235
Db 176 QLKMVTVKLISHREKQPHYGEVTTKMLCAADPQWKVTSQCGDGGPLVCSLQGRMTL 235
```

```
Db 553 SLREALVPLVADHKCSSPEVYGDADISPNMLCAGYFDCKSDACQDGGGLACEKNGVAYL 612
Qy 236 TCIVSWGRGCALKDKPGYVYTRVSHPELPIRSHTK 269
Db 613 YGIISWGDGCGRLHKGPGYVYTRVANYVDWINDRIR 646

RESULT 19
FA12_CAVPO STANDARD; PRT; 603 AA.
ID FA12_CAVPO
AC Q04962;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Butheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RX MEDLINE=93003367; PubMed=1390917;
RA Kambara T., Okabe H.;
RT "Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule.";
RL Biochim. Biophys. Acta 1159:113-121 (1992).
CC !- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC !- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla
CC !- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC !- SIMILARITY: Contains 1 fibronectin type I domain.
CC !- SIMILARITY: Contains 1 fibronectin type II domain.
CC !- SIMILARITY: Contains 2 EGF-like domains.
CC !- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68615; CAA48600.1;
DR PIR; S28941; S28941.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000883; FibrinctnI.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
```

DR PRINTS; PRO0013; ENTPBELL.
 DR PRINTS; PRO0018; KRINGLE.
 DR PRODOM; PD000995; FN_Type_II; 1.
 DR PRODOM; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN 1; 1.
 DR PROSITE; PS00023; FIBRONECTIN 2; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER . 1 1
 FT SIGNAL <1 18
 FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.
 FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.
 FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
 FT DOMAIN 93 130 EGF-LIKE 1.
 FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
 FT DOMAIN 173 209 EGF-LIKE 2.
 FT DOMAIN 216 294 KRINGLE.
 FT DOMAIN 312 342 PRO-RICH.
 FT DOMAIN 359 603 SERINE PROTEASE.
 FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 97 109 BY SIMILARITY.
 FT DISULFID 103 118 BY SIMILARITY.
 FT DISULFID 120 129 BY SIMILARITY.
 FT DISULFID 134 162 BY SIMILARITY.
 FT DISULFID 160 169 BY SIMILARITY.
 FT DISULFID 177 188 BY SIMILARITY.
 FT DISULFID 182 197 BY SIMILARITY.
 FT DISULFID 199 208 BY SIMILARITY.
 FT DISULFID 216 294 BY SIMILARITY.
 FT DISULFID 237 276 BY SIMILARITY.
 FT DISULFID 265 289 BY SIMILARITY.
 FT DISULFID 345 472 BY SIMILARITY.
 FT DISULFID 383 399 BY SIMILARITY.
 FT DISULFID 391 461 BY SIMILARITY.
 FT DISULFID 422 425 BY SIMILARITY.
 FT DISULFID 468 557 BY SIMILARITY.
 FT DISULFID 520 536 BY SIMILARITY.
 FT DISULFID 547 578 BY SIMILARITY.
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 503 AA; 56795 MW; 48DC6B946B9ED59 CRC64;

Query Match 32.2%; Score 485.5; DB 1; Length 603;
 Best Local Similarity 40.7%; Pred. No. 2.8e-36;
 Matches 111; Conservative 41; Mismatches 102; Indels 19; Gaps 8;

QY 6 PEELFQCCKQLRPRF-----KIGSEFTIENQWFAIYERHGGSVYVCGSLSP 61
 DB 338 PFTSSLCCQR-LRKLSSLRVGLVALPOAHPIYALY-----WGS--NFCSSLIAP 390
 QY 62 CWVISATHCFIDYPKKEDYVVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDI 121
 DB 391 CNVLTAHCLQNRAPEELKVVQLGQDRHQSCHEHQTQVLAHVSRLHEAFSPS--SYLNDL 448
 QY 122 ALLKI-RSKEGCAQPSRIOTICLPSMNDYDFG--TSCETIGGKENSIDYLPBOLK 178
 DB 449 ALLRLQKSDGSCAQLSPYVQVTCPLSPAPPSESETTCCEVAGHGQFEGAEYSFELQ 508

QY 179 MTVVKLISHRCQQPHYYGVSEVTTKMLCAADPMQWTDSCQDSCGGLVLC---SLOGRMTL 235
 DB 509 EAQPLISSERCSPSEVHGDAFLSMLCAGFLEGTDACQSDSGGLVCEDEAAHRLLIL 568
 QY 236 TGIVSWGRGCKALKKPGVYTVRVSHFLPWRSHST 268
 DB 569 RGIVSWGSGCGDRNKPQVYTDVASVLTWIKHT 601

RESULT 20
 ID FA12 BOVIN STANDARD; PRT; 593 AA.
 AC P98140;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAF) (Fragment).
 GN F12.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver.
 RX MEDLINE=94242782; PubMed=8186251;
 RA Shibuya Y., Semba U., Okabe H., Kambara T., Yamamoto T.;
 RT "Primary structure of bovine Hageman factor (blood coagulation factor
 RT XII): comparison with human and guinea pig molecules.";
 RL Biochim. Biophys. Acta 1206:63-70(1994).
 RN [2]
 RP SEQUENCE OF 10-21; 350-364 AND 525-550.
 RX MEDLINE=7182112; PubMed=861210;
 RA Fujikawa K., Walsh A.K., Davie W.E.;
 RT "Isolation and characterization of bovine factor XII (Hageman
 RT factor).";
 RL Biochemistry 16:2270-2278(1977).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor XIa.
 CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR
 CC XIIa ACTIVATES FACTOR XI TO FACTOR XIa. BOVINE FACTOR XII IS
 CC CLEAVED ONLY TO ALPHA-FACTOR XIIa AS IT LACKS THE TRYPSIN/
 CC KALLIKREIN CLEAVAGE SITE.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: S70164; AB330804.2; --
 CC PIR: S45281; S45281.
 CC HSRP; P00763; LDPO.
 CC MEROPS; S01.211; --
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000083; Fibinctnl.

DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Iry.
 DR Pfam; PF00008; EGF_2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN Type II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SMC0181; EGF; 2.
 DR SMART; SMC0058; FN1; 1.
 DR SMART; SMC0059; FN2; 1.
 DR SMART; SMC0130; KR; 1.
 DR SMART; SMC0020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; FALSE NEG.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 9
 FT CHAIN 10 349
 FT CHAIN 350 593
 FT CHAIN 37 78
 FT DOMAIN 84 121
 FT DOMAIN 123 163
 FT DOMAIN 164 200
 FT DOMAIN 207 287
 FT DOMAIN 297 333
 FT DOMAIN 350 593
 FT ACT_SITE 389 389
 FT ACT_SITE 438 438
 FT ACT_SITE 541 541
 FT ACT_SITE 88 100
 FT DISULFID 94 109
 FT DISULFID 111 120
 FT DISULFID 125 153
 FT DISULFID 151 160
 FT DISULFID 168 179
 FT DISULFID 173 188
 FT DISULFID 190 199
 FT DISULFID 207 287
 FT DISULFID 230 269
 FT DISULFID 258 282
 FT DISULFID 336 463
 FT DISULFID 374 390
 FT DISULFID 382 452
 FT DISULFID 413 416
 FT DISULFID 479 547
 FT DISULFID 510 526
 FT DISULFID 537 568
 FT CARBOHYD 99 99
 FT CARBOHYD 241 241
 FT CARBOHYD 263 263
 FT CARBOHYD 410 410
 SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;

Query Match 30.0%; Score 453; DB 1; Length 593;
 Best Local Similarity 38.5%; Pred No. 2.4e-33;
 Matches 102; Conservative 44; Mismatches 101; Indels 18; Gaps 7;
 QY 13 CGQ---KTLRPRKTIIGGETTIENQFWAAIYRHRGGSVTVVCGSLISPCWISATH 69

DB 336 CGQLRKWLSSLNRVVGGLVALPGAHPIYALYWDQ-----HFCAGSLIAPCWVLTAAH 389
 QY 70 CFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIR-S 128
 DB 390 CLQNRPAKELTVLQDQRHNSQCEQCQTAVRDYRLHEAFSPITYQH--DLALVRLQES 447
 QY 129 KEGRCQAQPSRTIQTICLPSMYNDPQFGTS--CEITGFGKENSTDYLYPEQLKMTVVKLIS 186
 DB 448 ADGCCAHPSPFVQVCLPSLTAARPAESAAYCEVAGMHQEGGE-YSSFLQEAQVPLID 506
 QY 187 HRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDSGGGLVC---SLQGRMTLTGTIVSMGR 243
 DB 507 PQRCSAPDVHGAATAQGLCAGFLEGGTDACQDSGGGLVCEDETPERQLILRGIVSNGS 566
 QY 244 GCALKDKPVGTVTRVSHFLPWIRSH 268
 DB 567 GCGNRLKEGVYTDVANYLAWIREHT 591
 RESULT 21
 NEUTR_HUMAN
 ID NEUTR_HUMAN STANDARD; PRT; 875 AA.
 AC P56730; Q9UP16;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurotrophin precursor (EC 3.4.21.-) (Mocopsin) (Leydin).
 GN PRSS12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98201705; PubMed=9540828;
 RA Proba K., Gschwend T.P., Sonderegger P.;
 RT "Cloning and sequencing of the cDNA encoding human neurotrophin.";
 RL Biochim. Biophys. Acta 1396:143-147 (1998).
 RV [2]
 R. SEQUENCE OF 615-875 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=99203523; PubMed=10103056;
 RA Poorafshar M., Hellman L.;
 RT "Cloning and structural analysis of leydin, a novel human serine
 protease expressed by the Leydig cells of the testis.";
 RL Eur. J. Biochem. 261:244-250 (1999).
 CC -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
 ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
 LEARNING AND MEMORY OPERATIONS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Brain and Leydig cells of the testis.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 4 SRCR domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ001531; CAA04816.1; -
 DR EMBL; AF077298; AAD25919.1; -
 DR HSSP; P00763; 1DPO.
 DR Genew; HGNC:9477; PRSS12.
 DR MIM; 606709; -
 DR MEROPS; S01.237; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000001; Kringle.


```

DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00051; kringle_1.
DR Pfam; PF00530; SRCR_4.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00258; SPERACTRCPTR.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00202; SR; 4.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS00420; SRCR_1; 3.
DR PROSITE; PS50287; SRCR_2; 3.
DR PROSITE; PS50287; SRCR_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 875 NEUTRYPSIN.
FT DOMAIN 23 92 PRO-RICH.
FT DOMAIN 93 165 KRINGLE.
FT DOMAIN 170 271 SRCR 1.
FT DOMAIN 280 381 SRCR 2.
FT DOMAIN 387 487 SRCR 3.
FT DOMAIN 500 601 SRCR 4.
FT DOMAIN 619 875 SERINE PROTEASE.
FT DOMAIN 619 630 ZWOGEN ACTIVATION REGION.
FT ACT_SITE 630 631 REACTIVE BOND (POTENTIAL).
FT ACT_SITE 676 676 CHARGE RELAY SYSTEM.
FT ACT_SITE 726 726 CHARGE RELAY SYSTEM.
FT ACT_SITE 825 825 CHARGE RELAY SYSTEM.
FT DISULFID 619 750 POTENTIAL.
FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 683 683 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 663 663 A -> V (IN REF. 2).
FT CONFLICT 701 701 E -> V (IN REF. 2).
FT CONFLICT 839 841 VVY -> AAL (IN REF. 2).
SQ SEQUENCE 875 AA; 97011 MW; B66EC946DC208DC8 CRC64;

Query Match 29.3%; Score 442.5; DB 1; Length 875;
Best Local Similarity 36.8%; Pred. No. 3.5e-32;
Matches 100; Conservative 47; Mismatches 113; Indels 13; Gaps 8;

QY 3 SGPPEBLKQCQGGKTLRPFK-IIGGEFTTIENQPMFAAIYRRHGGSVTVYVCGGSLISP 61
Db 609 NSNKESLSSVCGLLRHQKRIIGKNSLRGGWPQVSLRLKSSHGDRLLCGATLLSS 668
QY 62 CWVISATHCFIDYPKK-EDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db 669 CWVLTAAHCFKRYGNSRYSYAVRVGDYHTLVPEEFEEIGVQQIVIHREYRPRDSY-D 726
QY 121 IALLKTRSGEGRCAQPSRTIQTICLPSMYNDQ-FQTSCEITFGKENSTDVLYPEQLKM 179
Db 727 IALVRLQEEQCARSSSHVLPACLLWERQKTASNCYIIGW---DTGRAYRTIQQ 783
QY 180 TVVKLISHRECOQPHYGVSEVTTKMLCAAD--PQWTKDSQGDGSGPLVCSLQGR-MTLT 236
Db 784 AAIPLLPKFCEB--RYKGRFTGRMLCAGNLHKKRVDSQGDGSGPLMCPRGESVVVY 841
QY 237 GIVSMGRGALKDKPGVYTVRVSHLPWIBSHTK 269
Db 842 GVTSGYGGVGDSPGVITKVSFAFPWIKSVTK 874

RESULT 22
ID NEUTR_MOUSE
AC 008762;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Neurotrypsin precursor (EC 3.4.21.-) (Motopsin) (Brain-specific serine protease 3) (BSSP-3).

GN PRS12 OR BSSP3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]_TaxID=10090;

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=97401523; PubMed=9245503;

RA Gschwend T.P., Krueger S.R., Kozlov S.V., Wolter D.P., Sonderegger P.; "Neurotrypsin, a novel multidomain serine protease expressed in the nervous system.";

RT Mol. Cell. Neurosci. 9:207-219(1997).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98008848; PubMed=9344839;

RA Yamamura Y., Yamashiro K., Tsurucka N., Nakazato H., Tsujimura A., Yanaguchi N.;

RT "Molecular cloning of a novel brain-specific serine protease with a kringle-like structure and three scavenger receptor cysteine-rich motifs.";

RL Biochem. Biophys. Res. Commun. 239:386-392(1997).

CC -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH LEARNING AND MEMORY OPERATIONS.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: MOST ABUNDANT IN CEREBRAL CORTEX, HIPPOCAMPUS AND AMYGDALA.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: Contains 1 kringle domain.

CC -!- SIMILARITY: Contains 3 SRCR domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; Y13192; CAA73646.1; -.

DR EMBL; D89871; BAA23986.1; -.

DR HSSP; P00763; IDPO.

DR MEROPS; S01-237; -.

DR MGP; MGI:110881; Prss12.

DR InterPro; IPR00314; Chymotrypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser_protease_Try.

DR InterPro; IPR001190; Srcr_receptor.

DR Pfam; PF00051; kringle_1.

DR Pfam; PF00530; SRCR_3.

DR Pfam; PF00089; trypsin_1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00258; SPERACTRCPTR.

DR ProDom; PD000395; Kringle_1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00202; SR; 3.

DR SMART; SM00020; TRYPSIN_SPC; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS50070; KRINGLE_2; 1.

DR PROSITE; PS00420; SRCR_1; 3.

DR PROSITE; PS50287; SRCR_2; 3.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.

DR PROSITE; PS00135; TRYPSIN_SER; 1.

KW Hydrolyase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 761 NEUTRYPSIN.

FT DOMAIN 85 157 KRINGLE.

FT DOMAIN 166 267 SRCR 1.

FT DOMAIN 273 373 SRCR 2.

FT DOMAIN 386 487 SRCR 3.
 FT DOMAIN 505 761 SERINE PROTEASE.
 FT DOMAIN 505 516 ZMOGEN ACTIVATION REGION.
 FT ACT_SITE 516 517 REACTIVE BOND (POTENTIAL).
 FT ACT_SITE 562 562 CHARGE RELAY SYSTEM.
 FT ACT_SITE 612 612 CHARGE RELAY SYSTEM.
 FT ACT_SITE 711 711 CHARGE RELAY SYSTEM.
 FT DISULFID 505 636 POTENTIAL.
 FT CARBOHYD 93 93 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 521 521 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 569 569 N-LINKED (GLNAC. . .) (POTENTIAL).
 SQ SEQUENCE 761 AA; DF507B0371216436 CRC64;
 Query Match 28.28; Score 425.5; DB 1; Length 761;
 Best Local Similarity 37.14; Pred. No. 1e-30;
 Matches 101; Conservative 42; Mismatches 116; Indels 13; Gaps 8;
 QY 3 SSPPEELKFCGQKTLRPFK-IIGFEFTIENQPMFAAIYRHRGGSVTVYVGGSLISP 61
 DB 495 SGNKEMLSGGLRLHRRQKRIIGGNSLRGAWPMQASLRASAHGDRLLCGATLLSS 554
 QY 62 CWTISATHCIDY-PKEDYIVYLGSRNSNTQGMKFEVENLIHKDYSADTLAHND 120
 DB 555 CWTIAHCFKRYGNNSRSYAVRGVYHTLVPEFEQIGVQIVHRYRPRDSY--D 612
 QY 121 IALLKIRSEKRCACQAPERTTQICLPSMYNDPO-FGTSCBITGFKENSTDYLYPEQLKM 179
 DB 613 IALVRLOGPEQCARLSTHVLPAQLWRPPOKTSNCHITCWG--DTGRAYSTLQ 669
 QY 180 TVVKLISHRCQPHYGVSEVTTKMLCAADPW--KTDSCQSGSGPLVCSLOGR-MTIT 236
 DB 670 AAVPELLPRFCKE--RYKGLFTGRMLCAGNLQEDNRVDSQSGSGPLMCKEKPDESVMVY 727
 QY 237 GIVSWGSGCALDKPGVYTVRVSHFLPWRSH 268
 DB 728 CVTSWGYGCGVKUTPGVYTVRPAFVPIKSVT 759

RESULT 23

HATT HUMAN
 ID HATT HUMAN STANDARD; PRT; 418 AA.
 AC O60235;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Airway trypsin-like protease precursor (EC 3.4.21.-).
 GN HAT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98234382; PubMed=9565616;
 RA Yamaoka K., Masuda K.-I., Ogawa H., Takagi K., Umamoto N., Yamaoka S.;
 RT "Cloning and characterization of the cDNA for human airway trypsin-
 RT like protease.";
 RL J. Biol. Chem. 273:11895-11901(1998).
 RN [2]
 RN SEQUENCE OF 187-206, AND CHARACTERIZATION.
 RP MEDLINE=97224034; PubMed=9070615;
 RA Yamaoka S., Ohnishi T., Kawano S., Tsuchihashi S., Ogawara M.,
 RA Masuda K.-I., Yamaoka K., Takahashi M., Sano T.;
 RT "Purification, characterization, and localization of a novel
 RT trypsin-like protease found in the human airway.";
 RL Am. J. Respir. Cell Mol. Biol. 16:300-308(1997).
 CC -1- FUNCTION: May play some biological role in the host defense system
 CC on the mucous membrane independently of or in cooperation with
 CC other substances in airway mucous or bronchial secretions
 CC -1- CATALYTIC ACTIVITY: Preferentially cleaves the C-terminal side of
 CC arginine residues at the P1 position of certain peptides, cleaving
 CC Boc-Phe-Ser-Arg-4-methylcoumarinyl-7-amide most efficiently and
 CC having an optimum pH of 8.6 with this substrate.

CC -1- ENZYME REGULATION: Strongly inhibited by diisopropyl
 CC fluorophosphate, leupeptin, antipain, aprotinin, and soybean
 CC trypsin inhibitor, but hardly inhibited by secretory leukocyte
 CC protease inhibitor at 10 microm.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ACTIVATED BY
 CC CLEAVAGE AND SECRETED.
 CC -1- TISSUE SPECIFICITY: Located in the cells of the submucosal serous
 CC glands of the bronchi and trachea.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 1 SEA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB002134; BAA28691.1; --
 CC HSP; F00750; IRTF.
 CC MIM; 605369; --
 CC GO; GO:0005576; C:extracellular; TAS.
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0008233; P:peptidase activity; TAS.
 CC GO; GO:0007585; P:respiratory gaseous exchange; TAS.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000082; SEA domain.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF01390; SEA; 1.
 CC PRINTS; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00200; SEA; 1.
 CC SMART; SM00200; Tryp_Spc; 1.
 CC PROSITE; PS50024; SEA; 1.
 CC PROSITE; PS50240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase, Serine protease, Transmembrane; Signal-anchor; Zymogen;
 KW Glycoprotein.
 FT CHAIN 1 186 AIRWAY TRYPSIN-LIKE PROTEASE, NON-
 FT CATALYTIC CHAIN.
 FT AIRWAY TRYPSIN-LIKE PROTEASE, CATALYTIC
 FT CHAIN 187 418 CHAIN.
 FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT DOMAIN 42 418 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 44 164 SEA.
 FT DOMAIN 187 417 SERINE PROTEASE.
 FT ACT_SITE 227 227 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 368 368 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 173 292 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 212 228 BY SIMILARITY.
 FT DISULFID 337 353 BY SIMILARITY.
 FT CARBOHYD 144 144 N-LINKED (GLNAC. . .) (POTENTIAL).
 SQ SEQUENCE 418 AA; 46263 MW; F4BCIDB020CFBBD0 CRC64;
 Query Match 28.11; Score 423.5; DB 1; Length 418;
 Best Local Similarity 36.41; Pred. No. 7.4e-31;
 Matches 90; Conservative 47; Mismatches 93; Indels 17; Gaps 6;
 QY 23 KIIGFEFTIENQPMFAAIYRHRGGSVTVYVGGSLISPCWVATHCIFDYPKEDYIV 82
 DB 186 RILGTEAEESWPMQVSLRN-----AHCGSLINMILTAHCFRNSNPRDWIA 240
 QY 83 YLGRSLNSNTQGMKFEVENLIHKDYSADTLAHNDIALKIRSEKRCACQAPSRITQT 142
 DB 241 TSG---ISTTFPKLRMRVRNLIHNNYKSA--HENDIALVRLNS-----VTFTKDIHS 290

QY 143 ICLPSWYNDPQGTSCITGFKENSTDYLPKQMTWVLLSHRECCQPHYGVSEVTT 202
 DB 291 VCLPATQNIIPGTSAYTGTGWAQYAGHTYPE-LRQGVRIIENDVCNAPHSNGAILS 349
 QY 203 KMLCAADPQWTKDSQCGSGPLVCSLQGRM-TLTGIVSWGRGKALDKPKGVYTRVSHFL 261
 DB 350 GMLCAGVPQGGVDACQSGSGPLVQEDSRRLFWIVGIVSGDQCGLPKPKGVYTRVTAYL 409
 QY 262 PWIRSH 268
 DB 410 DWIROQT 416

RESULT 24
 PSS8 HUMAN
 ID PSS8 HUMAN STANDARD; PRT; 343 AA.
 AC Q16651; QUCU3;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Prostatein precursor (EC 3.4.21.-).
 GN PRS8.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Prostate;
 RX MEDLINE=95286644; PubMed=7768952;
 RA Yu J.X., Chao L., Chao J.;
 RT "Molecular cloning, tissue-specific expression, and cellular
 RT localization of human prostatein mRNA."
 RL J. Biol. Chem. 270:13483-13489 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalak U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE OF 45-64.
 RC TISSUE=Semen;
 RX MEDLINE=94308140; PubMed=8034638;
 RA Yu J.X., Chao L., Chao J.;
 RT "Prostatein is a novel human serine proteinase from seminal fluid.
 RT Purification, tissue distribution, and localization in prostate
 RT gland.";
 RL J. Biol. Chem. 269:18843-18848 (1994).
 RC -1- FUNCTION: POSSESSES A TRYPSIN-LIKE CLEAVAGE SPECIFICITY.
 RC -1- SUBUNIT: HETERODIMER OF TWO CHAINS, LIGHT AND HEAVY, HELD BY A
 RC DISULFIDE BOND.
 RC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF
 ITS C-TERMINUS.

CC -1- TISSUE SPECIFICITY: FOUND IN PROSTATE, LIVER, SALIVARY GLAND,
 CC KIDNEY, LUNG, PANCREAS, COLON, BRONCHUS AND RENAL PROXIMAL TUBULAR
 CC CELLS. IN THE PROSTATE GLAND IT MAY BE SYNTHESIZED IN EPITHELIAL
 CC CELLS, SECRETED INTO THE DUCTS, AND EXCRETED INTO THE SEMINAL
 CC FLUID.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L41351; AAC41759.1; -;
 CC EMBL; U33446; AAB19071.1; -;
 CC EMBL; BC001462; AAB01462.1; -;
 CC PIR; A57014; A57014.
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.159; -;
 CC Genew; HGNC:9491; PRS8.
 CC MIM; 600823; -;
 CC GO; GO:0005615; C:extracellular space; TAS.
 CC GO; GO:0005886; C:plasma membrane; TAS.
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC Pfam; PF00089; trypsin; 1
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; Tryp_SPC; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;
 KW Transmembrane.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 32 ACTIVATION PEPTIDE.
 FT CHAIN 33 44 PROSTATIN LIGHT CHAIN.
 FT CHAIN 45 322 PROSTATIN HEAVY CHAIN.
 FT PROPEP 323 343 POTENTIAL.
 FT TRANSMEM 320 340 SERINE PROTEASE.
 FT DOMAIN 45 286 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 37 154 BY SIMILARITY.
 FT DISULFID 70 86 BY SIMILARITY.
 FT DISULFID 168 244 BY SIMILARITY.
 FT DISULFID 201 223 BY SIMILARITY.
 FT DISULFID 234 262 BY SIMILARITY.
 FT ACT_SITE 85 85 CHARGE RELAY SYSTEM.
 FT ACT_SITE 134 134 CHARGE RELAY SYSTEM.
 FT ACT_SITE 238 238 CHARGE RELAY SYSTEM.
 FT CARBOHYD 159 159 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 343 AA; 36431 MW; 98DD6447F5A8C1B2 CRC64;
 Query Match 27.1%; Score 408.5; DB 1; Length 343;
 Best Local Similarity 36.6%; Pred. No. 1.3e-29;
 Matches 100; Conservative 36; Mismatches 100; Indels 37; Gaps 8;
 QY 13 CQKTLRPRFKIIGGEFTTIEQNPWFAIYRRHGGSVTY----VCGGSLISPCWVISA 68
 DB 37 CG---VAQARITGSSAVAGQWPHQV-----SITYEGVHVCGLSYEQWVLSAA 84
 QY 59 HCFIDYPKKEDYIVYGLSRNSNTQGMKEFVENLIHKDYSADTLAHNDIALLKTRS 128
 DB 85 HCFPEHHEKEAYEVKLGAHQLDSDYSDAKVSTLKDIIHPHSYLOE--GSGDIALQL-- 140
 QY 129 KEGRCAQPSRTIQTICLPSPMYNDPQGTSCITPGFK-ENSTDYLYPEQLKMTVVKLISH 187
 DB 141 --SRPITSRVIRPICLPAANASPNGLHCTVTGHWAPSVLLTPKPLQOLEVPLISR 198
 QY 188 REC-----QPHYGVSEVTTKMLCAADPQWTKDSQCGSGGGLVCSLQGRMTLTG 237
 DB 199 ETCNCLYNIDAKPEEPHF----VQEDMYCAGYVEGKDACQCGSGGGLSPVCEGLWYLTG 254

Qy 238 IVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 270
Db 255 IVSWGDACARRPGVYTLASSYASWISQSKVTE 287

RESULT 25

KAL_MOUSE
ID KAL_MOUSE STANDARD; PRT; 638 AA.
AC P26262;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (Kininogenin) (Fletcher factor).
GN K1KB1 OR K1K3 OR PK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BAUB/C; TISSUE=Liver;
RX MEDLINE=91090844; PubMed=2264928;
RA Seidah N.G., Sawyer N., Hamelin J., Milon P., Beaubien G.,
RA Brachpapa L., Rochemont J., McKay M., Chretien M.
RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,
RT and comparison of protein and mRNA levels among species."
RL DNA Cell Biol. 9:737-748(1990).
CC -!- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XIIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 4 apple domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M58568; AAA63393.1; -
CC PIR: A36557; KMSPL.
CC HSRF: P00750; IRTF.
CC MEROPS: S01.212; -
CC MGD: MGI:102849; K1kb1.
CC InterPro: IPR000177; Apple.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR003014; PAN.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam: PF00024; PAN; 4.
CC Pfam: PF00069; trypsin; 1.
CC PRINTS: PR00005; APPLEDOMAIN.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00223; APPLE; 4.
CC SMART: SM00020; Tryp_SPC; 1.
CC PROSITE: PS00495; APPLE; 4.
CC PROSITE: PS00240; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;

Repeat. 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 O-LINKED (PROBABLE).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (PROBABLE).
FT ACT_SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 104 BY SIMILARITY.
FT DISULFID 47 77 BY SIMILARITY.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 111 194 BY SIMILARITY.
FT DISULFID 137 166 BY SIMILARITY.
FT DISULFID 141 147 BY SIMILARITY.
FT DISULFID 201 284 BY SIMILARITY.
FT DISULFID 227 256 BY SIMILARITY.
FT DISULFID 231 237 BY SIMILARITY.
FT DISULFID 292 375 BY SIMILARITY.
FT DISULFID 318 347 BY SIMILARITY.
FT DISULFID 322 328 BY SIMILARITY.
FT DISULFID 340 345 BY SIMILARITY.
FT DISULFID 383 503 BY SIMILARITY.
FT DISULFID 419 435 BY SIMILARITY.
FT DISULFID 517 584 BY SIMILARITY.
FT DISULFID 548 563 BY SIMILARITY.
FT DISULFID 574 602 BY SIMILARITY.
SQ SEQUENCE 638 AA; 71368 MW; CC27C93AB1086599 CRC64;
Query Match 27.0%; Score 407; DB 1; Length 638;
Best Local Similarity 33.8%; Pred. No. 3.8e-29;
Matches 89; Conservative 55; Mismatches 103; Indels 16; Gaps 6;
Qy 17 TLPRPKITGGEFTTIENQFWFAIYRRHGGVTVVCGSLISPCWVISATHCFFIDYPK 76
Db 384 TTKINARIVGGTNGASLGEPWPQVSL--QVKLVSTQHLGCGSIIGROWLVLTAAHCFDGPY 441
Qy 77 KEDYIVVLGSRSLNSNTQGBMKFEVENLILHKDYSADTLAHHNDIALLLKIRSGRCAQP 136
Db 442 PDVWRIYGGILSLEITKETPSSRIKELIHQYKVS--EGNYDIALIKLQTP----LNY 495
Qy 137 SRTIQTICLPSMYNDPQFGTSCBITQFG---KENSTDYLYPEQLKMTVVKLISHRECQPP 193
Db 496 TEFQKPICLPSKADTTNITVNCWVTGWTGKEQGETONI---LQKATIPLVNNEECQK- 550
Qy 194 HYYGSEVTTKMLCAADPQWKTSCQDGSGLPVCISLQGRMTLTGIVSWGECALKDKPGV 253
Db 551 KYRDYVINKQMICAGYKEGGTDACKGDSGLPVCYKSGRWQLVGITSGEGCGRQDPGV 610
Qy 254 YTRVSHFLPWIRSHTKENGLAL 276
Db 511 YTKVSEYMDWILEKTQSSDVRL 633
RESULT 26
KAL_MOUSE
ID KAL_MOUSE STANDARD; PRT; 277 AA.
AC Q9UKR3; Q9Y433;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4)
DE (KLK-L4).
GN KLK13 OR KLK14.
OS Homo sapiens (Human).

Matches	101;	Conservative	34;	Mismatches	81;	Indels	30;	Gaps	10												
QY	26	QGEFTTIENQPF	AAIYRHRG	SVTVYCG	SLSPCW	WISATHC	FDIPK	KEDYIVYL	85												
				:	:	:	:														
Db	38	GGYTCTP	PHSPQ	AAALLVQ	CR-----	LLCG	GVLVHP	KVLTAAHCL	87												
				:	:	:	:														
QY	86	RSRLNSNTGEM	KFEVENL	LHKDY--	SADTLA	HNHDIAL	LKIRSK	EGRCAP	SRTIQT	143											
				:	:	:	:														
Db	88	KIALGRV	EAGEQ	RVVVHSI	PIPEY	RRSP	THLND	HIDML	LELQSP	142											
				:	:	:	:														
QY	144	CLP	SMYND---	PQFG	TSCHIT	FGK	ENST	DYIY	PEOL	KMTV	WV	KLIS	HR	EQ	PH	Y	G	SEV	200		
				:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	143	-LPLSH	NNRLTP-	GTT	CR	SV	GW	GT	TTS	PQ	VN	PK	T	LQ	C	A	N	I	Q	197	
				:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	201	TYTMLCA	ADP	QWK	TSC	GD	SG	PL	VC	SG	LR	Q	RT	L	G	I	V	SW	GR	259	
				:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	198	TDM	L	CAG	T	K	E	G	K	S	C	E	G	S	G	P	L	V	C	N	253
				:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
QY	260	FLP	WIR	265																	
				:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			
Db	254	YVL	WIR	259																	
				:	:	:	:	:	:	:	:	:	:	:	:	:	:	:			

RESULT 27

HEPS_MOUSE	STANDARD;	PRT;	436 AA.
ID	HEPS_MOUSE		
AC	Q354E3; Q9CWN97;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-SEP-2003 (Rel. 42, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Serine protease hepsin (EC 3.4.21.-)		
GN	HPN.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OC	NCBI_Taxid=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RP	TISSUE=Liver;		
RC	MEDLINE=98058912; PubMed=9395459;		
RA	Vu T.-K.H., Liu R.W., Haakema C., Tomaasek J.J., Howard E.W.;		
RT	"Identification and cloning of the membrane-associated serine		
RT	protease, hepsin, from mouse preimplantation embryos."		
RL	J. Biol. Chem. 272:31315-31320(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RP	MEDLINE=99339944; PubMed=10411637;		
RA	Kawamura S., Kurachi S., Devashiki Y., Kurachi K.;		
RT	"Complete nucleotide sequence, origin of isoform and functional		
RT	characterization of the mouse hepsin gene."		
RL	Eur. J. Biochem. 262:755-764(1999).		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RP	STRAN=C57BL/6J; TISSUE=Kidney;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsi G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohatsu S.,		
RA	Havashiz		

RA Suzuki H., Itoyo-oka K., Wang A.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Havaishizaki Y.

CC SUBFAMILY.
CC -!- SIMILARITY: Contains 4 apple domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13143; AAA60153.1; .
DR EMBL; AF232742; AAF79940.1; .
DR EMBL; AF232734; AAF79940.1; JOINED.
DR EMBL; AF232735; AAF79940.1; JOINED.
DR EMBL; AF232736; AAF79940.1; JOINED.
DR EMBL; AF232737; AAF79940.1; JOINED.
DR EMBL; AF232738; AAF79940.1; JOINED.
DR EMBL; AF232739; AAF79940.1; JOINED.
DR EMBL; AF232740; AAF79940.1; JOINED.
DR EMBL; AF232741; AAF79940.1; JOINED.
DR PIR; A00921; KQHP.
DR HSP; P00763; IDPO.
DR MEROPS; S01.212; .
DR Genew; HGNC:6371; KLB1.
DR MIM; 229000; .
DR GO; GO:0003807; F:plasma kallikrein activity; TAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR InterPro; IPR001177; Apple.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00005; APPLDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00223; APPLE; 4.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00495; APPLE; 4.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase, Serine protease, Glycoprotein, Plasma, Zymogen, Signal;
KW Fibrinolysis, Blood coagulation, Inflammatory response, Liver;
KW Repeat, Polymorphism.
FT SIGNAL 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .).
FT ACT SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT SITE 578 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 104
FT DISULFID 47 77
FT DISULFID 51 57
FT DISULFID 111 194
FT DISULFID 137 166
FT DISULFID 141 147
FT DISULFID 201 284
FT DISULFID 227 256
FT DISULFID 231 237
FT DISULFID 292 375
FT DISULFID 318 347
FT DISULFID 322 328

FT DISULFID 340 345
FT DISULFID 383 503
FT DISULFID 419 435
FT DISULFID 517 584
FT DISULFID 548 563
FT DISULFID 574 602
FT VARIANT 143 143 N -> S (COMMON POLYMORPHISM).
FT VARIANT 202 202 /FTId=VAR_013598.
FT VARIANT 208 208 H -> Q.
FT VARIANT 208 208 /FTId=VAR_013599.
FT VARIANT 208 208 H -> P.
FT SEQUENCE 638 AA; 71369 MW; B62F9C1053838FB4 CRC64; /FTId=VAR_013600.
Query Match 26.7%; Score 402; DB 1; Length 638;
Best Local Similarity 33.1%; Pred. No. 1.le-28;
Matches 88; Conservative 56; Mismatches 98; Indels 24; Gaps 7;
Qy 17 TLRPRKIIIGETTTIENOPWFAAIY-----RRHGGSVTVYCGGSLISPCWVISATHCF 71
Db 384 TTKTSIRIVGGTNSWGEPWQVSLQVKLTQAH-----LCGSLIGHQWVLTAAHCF 436
Qy 72 IDYPKEDYIVLGRSLNSNTQGENKFEVENILHKOYSADTLAHNDIALKIRKEG 131
Db 437 DGLPLQDQVWRIYSGILNLSDTXDPFSQIKETIIHQNYKVSEGNH--DIALIKLOAP-- 492
Qy 132 RCAQPSRTIQTICLPSMYNDPQFGTSCETIGF--KENSTDYLYPEQLKMTVVKLSHRE 189
Db 493 --LNYTEFQKPLCPKSGKGTSTIYTNCTWGWGFSKEKGE---IQNILQKNIPLVNTNEE 547
Qy 190 CQPHYVYGVSEVTTMLCAADPQWKTDSCQSDSGPLVCSLQGRMTLTIGIVSWGRCALKD 249
Db 548 CQR-RYQDYKITQRMVYKAGYKGGKACKGDSGGLVCKHNGWRLVGVITSWGEGCARRE 606
Qy 250 KPGVYTRVSHFLPWRSHTKENGLA 275
Db 607 QPGVYTKVAEYMDWILEKTQSSDGA 632
RESULT 29
TMSS MOUSE
ID -TMSS_MOUSE STANDARD; PRT; 455 AA.
AC Q9ER04; Q9ER02; Q9ER03;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).
GN TMPSR35.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RP TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "cDNA cloning of mouse spinesin."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A. (ISOFORM 4).
RP TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RT "Molecular cloning of mouse type 4 spinesin."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=4;
CC IsoId=Q9ER04-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9ER04-2; Sequence=VSP_005397; VSP_005398;
CC Name=2;
CC IsoId=Q9ER04-3; Sequence=VSP_005395;
CC

CC Name=3;
 CC TAOID-Q9ER04-4; Sequences=VSP_005396;
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SIMILARITY: Contains 1 SRCR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB016229; BAB20276.1; --
 CC EMBL; AB016230; BAB20277.1; --
 CC EMBL; AB016423; BAB20278.1; --
 CC EMBL; AB041037; BAB40328.1; --
 CC HSSP; P00763; 1DPO.
 CC MEROPS; S01.313; --
 CC MGD; MG1.1933407; Tmpres5.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC InterPro; IPR001190; Srcr_receptor.
 CC Pfam; PF00089; trypsin; 1.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; P500240; TRYP SIN DOM; 1.
 CC PROSITE; P500134; TRYP SIN HIS; 1.
 CC PROSITE; P500135; TRYP SIN SER; 1.
 CC PROSITE; P500420; SRCR_1; FALSE_NEG.
 CC PROSITE; P500287; SRCR_2; 1.
 CC Hydrolase; Serine protease; Transmembrane; Signal-anchor;
 KW Glycoprotein; Alternative splicing.
 FT DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 50 70 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT DOMAIN 71 455 EXTRACELLULAR (POTENTIAL)...
 FT DOMAIN 112 207 SRCR.
 FT DOMAIN 218 455 SERINE PROTEASE.
 FT ACT_SITE 258 258 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 308 308 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 405 405 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT SITE 217 218 CLEAVAGE (POTENTIAL).
 FT DISULFID 135 196 BY SIMILARITY.
 FT DISULFID 148 206 BY SIMILARITY.
 FT DISULFID 243 259 BY SIMILARITY.
 FT DISULFID 374 390 BY SIMILARITY.
 FT DISULFID 401 429 BY SIMILARITY.
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1 144 Missing (in isoform 2).
 FT VARSPLIC 1 10 Missing (in isoform 3).
 FT VARSPLIC 1 182 Missing (in isoform 1).
 FT VARSPLIC 183 192 Missing (in isoform 1).
 FT CONFLICT 325 325 GLVDAWAF -> MEAQGLLWV (in isoform 1).
 FT CONFLICT 325 325 D -> G (IN REF. 1; BAB20277).
 FT SEQUENCE 455 AA; 49632 MW; 5CF3C1789C6899AA CRC64;
 Query Match 26.4%; Score 398.5; DB 1; Length 455;
 Best Local Similarity 34.5%; Pred.No.1.5e-28;
 Matches 96; Conservative 42; Mismatches 99; Indels 41; Gaps 8;
 QY 2 P5PPEELK-FQCGQKTLPRFKIIIGETFTIENQFWFAAIIYRRHGGSVTVVCGGSLIS 60
 DB 197 PSGRIVSLKCECGARPLASR-IVGGQAVASGRFWQASVMLGSR-----HTCGASVLA 249
 QY 61 PCWVTSATHCFTDYPKEDYIYVLGSRSLNS-----NTQGMKFEVENIL 106
 DB 250 PHWVVTAAHCMTVSF-----RLSRSLWSRWVHAGLVSHGAVRQHQGTM---VEKIIP 296

QY 107 HKDYSADTLAHNDIALLLKIRSEKGRCAQPSRTIQTICLPSWYNDPQFTSCBITGFGKE 166
 DB 297 HPLYSAGN--HDYDVALQLRTP-----INFSDVDVAVCLPAKEQYPPWGSQCWVGWGT 350
 QY 167 NSTDYLYPEOLKMTVVKVLIHRESCQPHYVGSVVTMLCAADPOWKTDSCQDSDSGPLV 226
 DB 351 DPSTHSSDLDQTWPLLTSLHCNSCMYSGALTHRLMCAGYLDGRADACQDSDSGPLV 410
 QY 227 CSLOGRMTLFGIVSWGRCALKDKPGVYTVRSHFLPMI 264
 DB 411 CPSGDTWHLVGVVSWGRGCAEPNRPVGVYAKVAFLOWI 448

RESULT 30

HEPS_HUMAN
 ID HEPS_HUMAN STANDARD; PRT; 417 AA.
 AC P05981;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Serine protease hepsin (EC 3.4.21.-) (Transmembrane protease, serine
 DE 1).
 GN HPN OR TMPRSS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RX MEDLINE=88209431; PubMed=2835076;
 RA Levitus S.P., Loeb K.R., Hagen F.S., Kurachi K., Davie E.W.,
 RT "A novel trypsin-like serine protease (hepsin) with a putative
 RT transmembrane domain expressed by human liver and hepatoma cells.";
 RL Biochemistry 27:1067-1074 (1988).
 RN [2]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Pancreas, and Spleen;
 RX MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski W.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP CHARACTERIZATION
 RX MEDLINE=91358502; PubMed=1885621;
 RA Tsuji A., Torres-Rosado A., Aral T., le Beau M.M., Lemons R.S.,
 RA Chou S.H., Kurachi K.;
 RT "Hepsin, a cell membrane-associated protease. Characterization,
 RT tissue distribution, and gene localization.";
 RL J. Biol. Chem. 266:16948-16953 (1991).
 RN [4]
 RP CHARACTERIZATION
 RX MEDLINE=91348237; PubMed=8346233;
 RA Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
 RT "Hepsin, a putative cell-surface serine protease, is required for

mammalian cell growth.";
Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).
-!- FUNCTION: Plays an essential role in cell growth and maintenance
of cell morphology.
-!- SUBCELLULAR LOCATION: Type II membrane protein.
-!- TISSUE SPECIFICITY: Present in most tissues, with the highest
level in liver.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; M18930; AAA36013.1; --
EMBL; X07732; CAA30558.1; --
EMBL; X07002; CAA30058.1; --
EMBL; BC025716; AAR25716.1; --
PIR; S00845; S00845.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.224; --
DR Genew; HGNC:5155; HPN.
DR MIW; 142440; --
DR GO; GO:0005887; C: integral to plasma membrane; TAS.
DR GO; GO:0008151; P: cell growth and/or maintenance; TAS.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Transmembrane; Signal-anchor.
FT FT CHAIN 1 162
FT FT CHAIN 163 417
FT FT DOMAIN 1 17
FT FT TRANSMEM 18 44
FT FT
FT FT DOMAIN 45 417
FT FT DOMAIN 163 417
FT FT ACT_SITE 203 203
FT FT ACT_SITE 257 257
FT FT ACT_SITE 353 353
FT FT DISULFID 153 277
FT FT DISULFID 188 204
FT FT DISULFID 322 338
FT FT DISULFID 349 381
FT FT CARBOHYD 112 112
SQ SEQUENCE 417 AA; 45011 MW; B2086FF661E551D7 CRC64;
Query Match 26.3%; Score 396.5; DB 1; Length 417;
Best Local Similarity 35.8%; Pred.No.2e-28;
Matches 101; Conservative 51; Mismatches 89; Indels 41; Gaps 14;
Qy 13 CGQKTLPRFKIIGGEFTTIENQWFAAIYRRHRRGSGVTYVCGSLISPCWISATHCFI 72
Db 153 CGRRKL-PVDRIVGGRDTSLGRWPQVSL--RYDG---AHLGGSLISGDPVLTAAHCP- 205
Qy 73 DYPKKEDYIVYLGSRNSNTQGE-----MKFEVENILHKQY-----SADTLAHNDIAL 123
Db 206 --PERNR---VLSEWRVFAGAVQASPHGLQGLGVQAVVHGGYLPFRDPNSENSNDIAL 260
Qy 124 LKIRSEGRCAQPSRTIQTICPSMYNDPOFGTSCETIGKENSSTLYLPEQ---LKMT 180
Db 261 VHLSSP-----LPLTEYIQVCLPAGALVDGKICTVTGNG---NTQY-YGQAGVLQEA 312
Qy 181 VWKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGGLVC-----SLQGRMTLT 236

DB 313 RYPIISNDVNCGADFYGNQIKPKVFCAGYPEGGIDACQDGGPFVCEDSISRTPRWRLC 372
QY 237 GIVSWGRCALKDKPGVYTRVSHFLPW----IRSHTKEENGL 274
DB 373 GIVSWGTCGALAAQKPGVYTKVSDFREWIFQAIKTHS-EASGM 413

Search completed: January 8, 2004, 12:59:53
Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2004, 12:58:25 ; Search time 41 Seconds
(without alignments)
1737.135 Million cell updates/sec

Title: US-09-880-503-5
Perfect score: 1508
Sequence: 1 KPSPPPPELKQCCQKTLRP.....VSHPLPWIRSHTKENGLAL 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL23:

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.podent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.virus.*
- 16: sp.bacteriap.*
- 17: sp.archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1321	87.6	433	6 Q8MILO	Q8mi10 oryctolagus
2	1316	87.3	433	6 Q8MHY7	Q8mhy7 oryctolagus
3	670	44.4	214	6 Q9XT70	Q9xt70 oryctolagus
4	585	38.8	128	6 Q97587	Q97587 oryctolagus
5	583	38.7	516	4 Q8UJ99	Q8u99 homo sapien
6	573	38.0	395	4 Q8BZM1	Q8bz1 homo sapien
7	564.5	37.4	559	11 Q91VP2	Q91vp2 mus musculu
8	553	36.7	562	6 Q8SQ23	Q8sq23 sus scrofa
9	545.5	36.2	564	6 Q8MK31	Q8mk1 oryctolagus
10	497.5	33.0	653	11 Q8VCS4	Q8vcs4 mus musculu
11	495	32.8	103	6 Q95M99	Q95m9 equus cabal
12	495	32.8	300	4 Q96EF3	Q96ef3 homo sapien
13	492	32.6	615	4 Q91Z25	Q91z25 homo sapien
14	467	31.0	616	6 Q97507	Q97507 sus scrofa
15	463.5	30.7	868	5 Q9Y1V3	Q9y1v3 polyanthroca
16	451.5	29.9	597	11 Q35727	Q35727 mus musculu

17	442.5	29.3	327	4 Q8N171	Q8n171 homo sapien
18	435.5	28.9	284	4 Q8NF86	Q8nf86 homo sapien
19	426.5	28.3	267	5 Q9BK47	Q9bk47 luidia foli
20	425.5	28.2	505	5 Q966V4	Q966v4 halocynthia
21	425	28.2	517	11 Q8K0D2	Q8kod2 mus musculu
22	421.5	28.0	751	11 Q93JC8	Q93jc8 rattus norv
23	420	27.9	471	11 Q8CFE0	Q8cfe0 mus musculu
24	418	27.7	537	4 Q9BYE1	Q9bye1 homo sapien
25	417	27.7	560	4 Q14520	Q14520 homo sapien
26	414	27.5	581	4 Q9BYE2	Q9bye2 homo sapien
27	411.5	27.3	276	11 Q8CGR6	Q8cgr6 mus musculu
28	411	27.3	638	11 Q8R0P5	Q8r0p5 mus musculu
29	406.5	27.0	739	11 Q9DBI0	Q9dbi0 mus musculu
30	404.5	26.8	802	4 Q8IUE2	Q8iue2 homo sapien
31	404.5	26.8	811	4 Q8IU80	Q8iu80 homo sapien
32	402.5	26.7	371	11 Q8CJ16	Q8cj16 rattus norv
33	402.5	26.7	445	11 Q8CJ17	Q8cj17 rattus norv
34	396.5	26.3	455	11 Q8CDR0	Q8cdr0 mus musculu
35	394.5	26.2	277	5 Q96899	Q96899 scolopendra
36	389.5	25.8	453	11 Q8VDE0	Q8vde0 mus musculu
37	388.5	25.8	453	11 Q8KIT0	Q8kit0 mus musculu
38	384.5	25.5	339	11 Q95L44	Q95l44 mus musculu
39	384	25.5	340	11 Q8BJV6	Q8bjv6 mus musculu
40	383	25.4	263	11 Q9DC86	Q9dc86 mus musculu
41	382	25.3	263	11 Q9CR35	Q9cr35 mus musculu
42	382	25.3	624	11 Q9DAT3	Q9dat3 mus musculu
43	381	25.3	249	11 Q9QYN4	Q9qyn4 mus musculu
44	381	25.3	276	11 Q9QYN3	Q9qyn3 mus musculu
45	381	25.3	767	13 Q9DGR2	Q9dgr2 xenopus lae

ALIGNMENTS

RESULT 1

Q8MILO	PRELIMINARY;	PRT;	433 AA.
ID	Q8MILO		
AC	Q8MILO;		
DT	01-OCT-2002 (T-EMBLrel. 22, Created)		
DT	01-OCT-2002 (T-EMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (T-EMBLrel. 23, Last annotation update)		
DE	Urokinase-type plasminogen activator.		
GN	PLAU.		
OS	Oryctolagus cuniculus (Rabbit).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.		
OX	NCBI_TaxID=9986;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22155945; PubMed=12149463;		
RA	Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,		
RA	Dichek D.A.;		
RT	"Increased expression of urokinase during atherosclerotic lesion		
RT	development causes arterial constriction and lumen loss, and		
RT	accelerates lesion growth."		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.		
DR	EMBL; AV122285; AAM83187.1; -		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR006209; EGF_like.		
DR	InterPro; IPR000001; Kringle.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00051; kringle; 1.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PRO0722; CHYMOTRYPSIN.		
DR	PRINTS; PRO0015; KRINGLE.		
DR	ProDom; PD000395; Kringle; 1.		
DR	SMART; SM00130; KR; 1.		
DR	SMART; SM00020; Tryp_spc; 1.		
DR	PROSITE; PS00022; EGF 1; 1.		
DR	PROSITE; PS00021; KRINGLE 1; 1.		
DR	PROSITE; PS50070; KRINGLE_2; 1.		

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65664F36415549B0 CRC64;

Query Match 87.6%; Score 1321; DB 6; Length 433;
Best Local Similarity 87.3%; Pred. No. 8.9e-124; Indels 0; Gaps 0;
Matches 241; Conservative 15; Mismatches 20;

QY 1 KPSPPEELKFCQGQKTLRPFKIIIGGFTTIENQWFAALYRRHGGSVTVYVCGSLIS 60
DB 158 KPALPPGKLEFCQGQKALRPFKIIIGGFTTIENQWFAALYRRHGGSVTVYVCGSLIS 217
QY 61 PCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLLHKDYASADTLAHND 120
DB 218 PCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLLHKDYASADTLAHND 277
QY 121 IALLKIRSEKRCQAQPSRSIQICLPNNADPNFGTSCETITGFGKENSTDLYYPEQLKMT 180
DB 278 IALLKILSNNGCAQPSRSIQICLPNNADPNFGTSCETITGFGKENSTDLYYPEQLKMT 337
QY 181 VVKLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLOGRWTLTGIVS 240
DB 338 VVKLVSYQECQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLOGRWTLTGIVS 397
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
DB 398 WGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 433

RESULT 2

Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Matanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA, complete cds."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -;
DR EMBL; AB087224; BAC02685.1; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; K5; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.

DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 87.3%; Score 1316; DB 6; Length 433;
Best Local Similarity 87.0%; Pred. No. 2.8e-123; Indels 0; Gaps 0;
Matches 240; Conservative 15; Mismatches 21;

QY 1 KPSPPEELKFCQGQKTLRPFKIIIGGFTTIENQWFAALYRRHGGSVTVYVCGSLIS 60
DB 158 KPALPPGKLEFCQGQKALRPFKIIIGGFTTIENQWFAALYRRHGGSVTVYVCGSLIS 217
QY 61 PCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLLHKDYASADTLAHND 120
DB 218 PCWVISAHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLLHKDYASADTLAHND 277
QY 121 IALLKIRSEKRCQAQPSRSIQICLPNNADPNFGTSCETITGFGKENSTDLYYPEQLKMT 180
DB 278 IALLKILSNNGCAQPSRSIQICLPNNADPNFGTSCETITGFGKENSTDLYYPEQLKMT 337
QY 181 VVKLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLOGRWTLTGIVS 240
DB 338 VVKLVSYQECQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLOGRWTLTGIVS 397
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 276
DB 398 WGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 433

RESULT 3

Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J., Idell S.;
RL "Partial mRNA of rabbit uPA";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097847; AAD39351.1; -;
DR HSP; P00749; IEUN.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 44.4%; Score 670; DB 6; Length 214;
Best Local Similarity 84.2%; Pred. No. 6.1e-59;
Matches 123; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 KSSPPEELKFOCGQKTLRPRKIIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLSIS 60
DB 69 KPALPPKLEFQCGQKALRPRKIIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLSIS 128
QY 61 PCWVISATCFDYPKKEDEYVYVLRGRSLNSNTQGMKEFEVENLILHKDYSADTLAHND 120
DB 129 PCWVVSATCFINHQKEDYVYVLRGRSLNSNTQGMKEFEVEQLILHEGYRADTLAHND 188
QY 121 IALLKIRSKEGRCQAQPSRTIQICLP 146
DB 189 IALLKILSNNGCAQPSRSIQICLP 214

RESULT 4

Q97587 ID Q97587 PRELIMINARY; PRT; 128 AA.
AC Q97587; DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase (Fragment)
OS Oryctolagus cuniculus (Rabbit)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.D., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
RT and their potential regulators in the healing medial collateral
RT ligament."
RL Biochem. Biophys. Res. Commun. 252:757-763(1998).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF069711; AAC95003.1; --
DR HSSP; P00749; LEUN.
DR MEROPS; S01.231; --
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolyase; Kinase; Protease; Serine protease.
FT NON_TER 1 1
FT TER 128 128
SQ SEQUENCE 128 AA; 14328 MW; 1BC7ED30E071A06D CRC64;

Query Match 38.8%; Score 585; DB 6; Length 128;
Best Local Similarity 84.4%; Pred. No. 1e-50;
Matches 108; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 37 WFAAIYRRHRGGSVTVVCGSLSISPCWVISATCFDYPKKEDEYVYVLRGRSLNSNTQGE 96
DB 1 WFAAIYRRHRGGSVTVVCGSLSISPCWVISATCFINHQKEDYVYVLRGRSLNSNTQGE 60
QY 97 MKFEVENLILHKDYSADTLAHNDIALKIRSKEGRCQAQPSRTIQICLPMSYNDPQFGT 156
DB 61 MKFEVEQLILHEGYRADTLAHNDIALKILSNNGCAQPSRSIQICLPWNADPNEG 120
QY 157 SCEITGFG 164
DB 121 SCEITGFG 128

RESULT 5

Q9BU99 ID Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99; DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC002795; AAH02795.1; --
DR HSSP; P00750; IAS5H.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000195; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;
Query Match 38.7%; Score 583; DB 4; Length 516;
Best Local Similarity 44.9%; Pred. No. 9.9e-50;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-OKTLRPRFKIIGGETTIENQWPFAAIYRRH-RGGSVTVVCGSLSISPCWVISATHC 70
DB 253 CGLRQYQPPQRIKGGFLFADIASHPWQAAIFAKHRRSPGERFLCGGILISCWILSAHC 312
QY 71 FIDYPKKEDEYVYVLRGRSLNSNTQGMKEFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 313 FQERFPPHLLTVILGRTRYVVPGEEOKEFEKVIHKEFDDDT--YNDIALIQLKSDS 370
QY 131 GRCAQPSRTIQICLPMSYNDPQFGTSCITGKENSVDYLPQELKMTVVKLISHREC 190
DB 371 SRCQAESVTVVCLPPADQLPDWTECELSGKGKHEALSPFYSERLKEAHVRLYPSRC 430
QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQGGSGGFLVCSLQGRMTLTGIVSGRG 244
DB 431 TSQHLNRTVTDNMLCAGDTRSGQPQANLHDAQGGSGGFLVCLNDGRMTLVGIISGLG 490
QY 245 CALKDKPGVTRYSHFLPWIRSHTK 269
DB 491 CGQKDVPGVTVKVTNYLDWIRDNR 515
RESULT 6
Q9BZW1 ID Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1; DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE T-plasminogen activator.
 DS Sus scrofa (Pig).
 DC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 DC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 DX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Enamel organ;
 RA Ding Y., Xue J., Bartlett J.D.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; AF364605; AAM00297.1; -.
 DR HSP; P00761; IAN1.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01233; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 36.7%; Score 553; DB 6; Length 562;
 Best Local Similarity 41.9%; Pred. No. 1.le-46;
 Matches 111; Conservative 43; Mismatches 101; Indels 10; Gaps 4;

13 CGQKTLR-PRFKIIGGEFTTIENQPFAYI-RRHGGSVTVYVCGSLSPCWVISAHC 70
 299 CGLRQYKQPKQRIKGLYADITSHPWQAIFVKNRRSPGERFLCGGILISSCWVLSAHC 358

71 FIDYPKEDIVYVGRSLNSNTQGMKPEVENLILHKDYSADTLAHNDIALKIRSK 130
 359 FQRFPPHVRVVLGRTYRLVPGEEQAQVEVYIVHKEFDDDT--YNDIALQLKSDS 416

131 GRCAQPSRTIQTICLPSMYNDPOFGTSCETITGFKENSTDYLYPEQLKMTVVKLSHREC 190
 417 LTCQAESDAVTVCLPEANLQLPDWTCELSGYGKHEASPPYSERLKEAHRVLYPSSRC 476

191 QQPHYGVSEVTTKMLCAADPQW-----KTDSCGDSGGPLVCSLQGRMTLTGIVSWG 244
 477 TSKHLFNKTIINMLCAGDTRSGGANLHDACQDGGPLVCMKGNHMTLVGVISWGLG 536

245 CALKDKPGVYTRVSHFLPWIRSH 269
 537 CGQKDEGVYTKVTVNLNIRDNR 561

RESULT 9
 BMKBI PRELIMINARY; PRT; 564 AA.
 D Q8MKBI;
 C Q8MKBI;
 T 01-OCT-2002 (TrEMBLrel. 22, Created)
 T 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 T 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 E Tissue-type plasminogen activator.

OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; AY029518; AAK40240.1; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 36.2%; Score 545.5; DB 6; Length 564;
 Best Local Similarity 43.1%; Pred. No. 6.4e-46;
 Matches 115; Conservative 36; Mismatches 103; Indels 13; Gaps 4;

12 CQCKTLR-PRFKIIGGEFTTIENQPFAYI-RRHGGSVTVYVCGSLSPCWVIS 66
 297 QCATCGLRQDQPKQRIKGLFTDITAHFWQAIFVKNRRSPGERFLCGGILINSCWVLS 356

67 ATHCFIDYPKEDIVYVGRSLNSNTQGMKPEVENLILHKDYSADTLAHNDIALK 126
 357 AAHCFLERFPQKURVILGRYPLVSAEEQIFVEQILHERFDEGT--YNDIALKL 414

127 RSKEGRCAQPSRTIQTICLPSMYNDPOFGTSCETITGFKENSTDYLYPEQLKMTVVKLS 186
 415 KSTSGCAQESQAVRLVCLPDASLQLPDWTCELSGYGKHEEPSVFSQKAEHVRVLYP 474

187 HRCQPHYGVSEVTTKMLCAADPQW-----KTDSCGDSGGPLVCSLQGRMTLTGIVS 240
 475 SSRCTPQQLKRTVTGNMLCAGDTRSGGAQVNLHDACQDGGPLVCMTGDGHTLIGIIS 534

241 WGRGCAKDKPGVYTRVSHFLPWIRSH 267
 535 WGLCGGQKDEGVYTKVTVNLNIRNQ 561

RESULT 10
 Q8VCS4 PRELIMINARY; PRT; 653 AA.
 ID Q8VCS4
 AC Q8VCS4;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical 70.6 kDa protein.

RESULT 11
Q95M89 PRELIMINARY; PRT; 103 AA.
AC Q95M89;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator urokinase (Fragment).
GN PLAU.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21314992; PubMed=11421942;
RA Shubitowski D.M., Venta P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;
RT "Polymorphism identification within 50 equine gene-specific sequence
tagged sites".
RL Anim. Genet. 32:78-78(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY008806; AAK14840.1; JOINED.
DR EMBL; AY008803; AAK14840.1; JOINED.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 103
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;
Query Match 32.8%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 8.3e-42;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
QY 37 WFAAIYRRHGGSVTVYCGSLISPCWVLSATHCFIDYKEDYIVYLGSRSLNSNTQGE 96
DB 1 WFAAIYRRHGGSVTVYCGSLISPCWVLSATHCFINYPKEDYIVYLGSRSLNSNTQGE 60
QY 97 MKEVEENLILHKDYSADTLAHNDIALKIRSGRCAQPSRT 139
DB 61 MKEVEKLILHEDYSADTLAHNDIALKISSGTGCAQPSRS 103
RESULT 12
Q96EF3 PRELIMINARY; PRT; 300 AA.
AC Q96EF3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Strausberg R.;
RA Tissue=Lung;
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC012390; AAH12390.1; JOINED.
DR HSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.

Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; JOINED.
DR HSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR008209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00035; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYPRII.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000995; FN_Type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 653
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;
Query Match 33.0%; Score 497.5; DB 11; Length 653;
Best Local Similarity 38.5%; Pred. No. 4.9e-41;
Matches 105; Conservative 44; Mismatches 109; Indels 15; Gaps 6;
QY 2 PSSPPEELKFCQCKQ-----TLPRFKILGGEFTTIENQPHFAIYRRHGGSVTVYCGG 56
DB 382 PESAP-AVRPTCGKHKKRTLRPR--LIGSSSLPGSHPLAIIY----IGNS---FCAG 432
QY 57 SLISPCWISATHCFIDYKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLA 116
DB 433 SLVHTCWVSAAHCFANSPRDSITVLGQHFFNRTDVTQFGLKVPVYTLXSVFNPN 492
QY 117 HNDIALKIRSGRCAQPSRTITICLPMSYNDPQGTSCETGFGKENSSTYLYPEQ 176
DB 493 NH-DLVLRLLKKGRCVRSQFQFICLPEAGSSFPPTGHCQIAGWGHMDENSVSSNS 551
QY 177 LKQTVKLIHRECOQPHYVYGGSEVTTKMLCAADPOWKTDSQGSGLVCSLOGRMTLT 236
DB 552 LLEALVPLVADHKCSSPEVYAGDISPNMLCAGYFDCKSDACQSGSGGLVCEKNGVAYLY 611
QY 237 GIVSGRGCAKDKPGVYTRVSHFLWIRSHTK 269
DB 612 GIISGDDGGRNLKPGVYTRVANYVDWINDRIR 644

DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hypothetical protein; Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF382F636A CRC64;

Query Match 32.8%; Score 495; DB 4; Length 300;
 Best Local Similarity 38.5%; Pred. No. 3.3e-41;
 Matches 107; Conservative 44; Mismatches 109; Indels 18; Gaps 6;

QY 1 KPSSPPEELK---FQCGQ---KTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVC 54
 DB 29 KRQPSLIRNGSLSCQRLKSLSMTRVVGGLVALRGAHPVIAALYNGHS-----FC 82
 QY 55 GGLISPCWVISATHCFIDYPKKEDIVYVLRGRSLNSNTQGMKFEVENLIHKDYASDT 114
 DB 83 AGSLIAPCWLTAAHCLQDRPAPEDLTIVLQGBRRNHSCEPQOTLAVRSYRLHEAFS--P 140
 QY 115 LAHNDIALKIR-SKEGRCAQPSRTITCLPSMYNDPQFGTSCBITGFGKENSVDLY 173
 DB 141 VSYQHDALLRLQEDADGSCALLSPYQVCLPSGAARSETTLQCVAGMGHQFEGAEY 200
 QY 174 PEQLKMTVVKLIHRECCQPHYGVSEVTHMLCAADPQWKTDSCQDGGPLVCSLQ-- 231
 DB 201 ASFLQEAQVFLSLERCSAPDVHSGSILPGMLCAGFLGGTDACQDGGPLVCEQAAE 260
 QY 232 -RMTLTGIVSWGRCALKKPKGVTVVSHFLPWIRSH 268
 DB 261 RRLTLQGIISWGGCGDRNKGVTVDVAYYLAWIREHT 298

RESULT 13
 Q81225
 ID Q81225 PRELIMINARY; PRT; 615 AA.
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Coagulation factor XII-Mie.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wada H., Nishiooka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
 RT "Molecular characterization of coagulation factor XII-Mie";
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB095845; BAC23095.1; -
 DR EMBL; AB095845; BAC23095.1; -
 SQ SEQUENCE 615 AA; 67735 MW; 030508870AOC7EDB CRC64;

Query Match 32.8%; Score 492; DB 4; Length 615;
 Best Local Similarity 38.5%; Pred. No. 1.6e-40;
 Matches 107; Conservative 43; Mismatches 110; Indels 18; Gaps 6;

QY 1 KPSSPPEELK---FQCGQ---KTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVC 54
 DB 344 KRQPSLIRNGSLSCQRLKSLSMTRVVGGLVALRGAHPVIAALYNGHS-----FC 397
 QY 55 GGLISPCWVISATHCFIDYPKKEDIVYVLRGRSLNSNTQGMKFEVENLIHKDYASDT 114
 DB 398 AGSLIAPCWLTAAHCLQDRPAPEDLTIVLQGBRRNHSCEPQOTLAVRSYRLHEAFS--P 455
 QY 115 LAHNDIALKIR-SKEGRCAQPSRTITCLPSMYNDPQFGTSCBITGFGKENSVDLY 173
 DB 456 VSYQHDALLRLQEDADGSCALLSPYQVCLPSGAARSETTLQCVAGMGHQFEGAEY 515
 QY 174 PEQLKMTVVKLIHRECCQPHYGVSEVTHMLCAADPQWKTDSCQDGGPLVCSLQ-- 231
 DB 516 ASFLQEAQVFLSLERCSAPDVHSGSILPGMLCAGFLGGTDACQDGGPLVCEQAAE 575
 QY 232 -RMTLTGIVSWGRCALKKPKGVTVVSHFLPWIRSH 268

DB 576 RRLTLQGIISWGGCGDRNKGVTVDVAYYLAWIREHT 613

RESULT 14
 Q97507
 ID Q97507 PRELIMINARY; PRT; 616 AA.
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE FXII.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Takahashi T., Kihara T.;
 RT "Porcine liver factor XII";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL; AB022426; BAA37148.1; -
 DR HSSP; P00763; IDPO.
 DR MEROPS; S01.211; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000742; EGF_2
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00732; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00059; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 616 AA; 68012 MW; 4C5FB3D71EBBD1A9 CRC64;

Query Match 31.0%; Score 467; DB 6; Length 616;
 Best Local Similarity 37.8%; Pred. No. 5.2e-38;
 Matches 107; Conservative 46; Mismatches 102; Indels 28; Gaps 9;

QY 4 SPPEEL-----KFCQCKTLRPRF-----KIIGGEFTTIENQPFPAIYRRHRGGSVTV 52
 DB 342 APPQGRPLPSAGLVGGQQR-LRRLSSLNRIVGLVALPGAHPIIYALYNGQN----- 394
 QY 53 VCGSLISPCWVISATHCFIDYPKKEDIVYVLRGRSLNSNTQGMKFEVENLIHKDYSA 112
 DB 395 FCAGSLIAPCWLTAAHCLQNRPAPELTIVLQGBRRNHSCEPQOTLAVRSYRLHESYSP 454

QY 113 DTLAHNDIALKIR-SKEGRCAQSPRTTQTICLP---SWYNDPQGTSCETITGFGKENS 168
 DB 455 KTYCH--DLALVKETADCCAHSPFPVQVCLPRSVASSABPE-GALCEVAGWGHQPE 511
 QY 169 TDVLYPEQLQMTVKLISHRECOQPHYYSSEVTTKMLCAADPQWKTDSCQDGGGGLVC- 227
 DB 512 GABEYSSFLOEAQVPLISPERCSAADVHGAFTPGMLCAGFLEGGTACQDGGGGLVCE 571
 QY 228 --SIQGRMTITGIVSMGRGKALKDKPGVTVRVSHFLPWIRSH 268
 DB 572 DETAERGLVLRGIVSMGSGCGDLKPGVYTDVANYLAWIQEHT 614

RESULT 15
 ID QY1V3 PRELIMINARY; PRT; 868 AA.
 AC QY1V3;
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Tunicate retinoic acid-inducible modular protease precursor.
 GN TRAMP.
 OS Polyandrocarpa misakiensis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Styelidae; Polyandrocarpa.
 OX NCBI_TaxID=7723;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=white spot;
 RX MEDLINE=99423646; PubMed=10491255;
 RA Ohashi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
 RT "A retinoic acid-inducible modular protease in budding ascidians";
 RL Dev. Biol. 214:38-45(1999).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; AB030007; EMBL; X92571; CAA67891.1; -;
 DR HSP; P00763; IDPO.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR002172; LDL receptor_A.
 DR InterPro; IPR003609; Pan app.
 DR InterPro; IPR001254; Ser. protease Try.
 DR InterPro; IPR001190; Srcr. receptor.
 DR Pfam; PF00057; ldl_recept_a; 3.
 DR Pfam; PF00530; SRCR; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR PRINTS; PR00258; SPERACTRCPTR.
 DR SMART; SM00192; LDLa; 3.
 DR SMART; SM00473; PAN_AP; 1.
 DR SMART; SM00202; SR; 2.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS01209; LDLRA_1; 3.
 DR PROSITE; PS00068; LDLRA_2; 3.
 DR PROSITE; PS02087; SRCR_2; 2.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease; Signal.
 FT SIGNAL 28
 SQ SEQUENCE 868 AA; 97660 MW; F71462865F36A6CA CRC64;

Query Match 30.7%; Score 463.5; DB 5; Length 868;
 Best Local Similarity 38.9%; Pred. No. 1; Re-37;
 Matches 107; Conservative 44; Mismatches 103; Indels 21; Gaps 10;
 QY 2 PSSPPEELKFCQOKTL-----RPRFKIIGFETIENQPFALYRRHSGSVTVYCGGS 57
 DB 601 PTPPPMP---ECGRKPVIEAPLPTARIVGSGTEPHEWPMQAGIWL-----PTYWCGGS 652
 QY 58 LISPCWVISATHCFI-DYPKKEDIVVGLSRNSNTQGEWKEVENLILHKDYSADTLA 116
 DB 653 LIHPCWVLTAAHCFVREIPR-DYTIKGLDHITGVDDTEQLFKIAEIKH-DYNTVT-- 708

QY 117 HHNDIALKIRSKGRCAQSPRTTQTICLPMSYNDPQGTSCETITGFGKENS TDY-LYPE 175
 DB 709 KENDIALIRIENDARECATITPEVQTVCLPKSSQFPAKTICEVTGKGSATAVRAYVP 768
 QY 176 OLKMTVVKLISHRECOQPHYYSSEVTTKMLCAADPQWKTDSCQDGGGGLVCSLQG--RM 233
 DB 769 VLQEAIEPLANKKCLRSEY-TOLGPTMFCAGVLTGKDKSCQDGGGGLSCROSDDDRY 827
 QY 234 TLTGIVSMGRGKALKDKPGVTVRVSHFLPWIRSH 268
 DB 828 YVMGIVSMGNGCAKPKAGVYAKVAVFIDWIEQMT 862

RESULT 16
 ID O35727 PRELIMINARY; PRT; 597 AA.
 AC O35727;
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Factor XII.
 GN F12.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Schloesser M., Schwager S., Engel W.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR EMBL; X92571; CAA67891.1; -;
 DR HSP; P00760; IAQ7.
 DR MEROPS; S01.211; -;
 DR MGD; MGI:1891012; F12.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser. protease Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00133; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
 KW Serine protease.
 SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

```
Query Match 29.3%; Score 451.5; DB 11; Length 597;
Best Local Similarity 37.0%; Pred. No. 1.8e-36;
Matches 97; Conservative 47; Mismatches 103; Indels 15; Gaps 5;

QY 13 CGQ---KTLRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVYVCGGSLISPCWVISAATH 69
DB 341 CGQRFKGLSSFMVGVGLVALPGSHPIAALYGN---FCAGSLIAPCWVLTAAH 394

QY 70 CFIDYPKEDYIVYVGLGRSLNSNTQGENKFEVENLIHKDYSADTLAHNDIALKIR-S 128
DB 395 CLONRPAPEELTVVLGQDRHNCSEWCQTLAVRSYRLHEGFSSTYQH--DLALLRLQES 452

QY 129 KEGRCQAQSRITQICLPSMYNDPQFGTSCEITGFGKENS'DYLYPQLKMTVVKLISHR 188
DB 453 KNSCALISPHVQVCLPSGAAPSETVLCVAGNGHGLEABEYSYFLQEAQVFFALD 512

QY 189 ECQPHYYGVSEVTTKMLCAADPQWKTSCQDGGPLVC---SLQGRMTLIGVSWGRGC 245
DB 513 RCSNSNVHGDAILPMLCAGFLGGTDACQDGGPLVCEEGTAHQTLRGLVISWSGSC 572

QY 246 ALKDKPGVYTVSHFLPWIRSH 267
DB 573 GDRNKPQGYTVDVANYLAWIQXH 594

RESULT 17
ID Q8N171 PRELIMINARY; PRT; 327 AA.
AC Q8N171;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Similar to protease, serine, 8 (Prostasin) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; EC036846; AAH36846.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
FT NON TER 1
SQ SEQUENCE 327 AA; 35106 MW; 60458ED817AC1CF3 CRC64;

Query Match 29.3%; Score 442.5; DB 4; Length 327;
Best Local Similarity 38.8%; Pred. No. 6.6e-36;
Matches 104; Conservative 35; Mismatches 96; Indels 33; Gaps 8;

QY 13 CGQKTLRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVYVCGGSLISPCWVISAATHCFI 72
DB 75 CGQPRMSSR--IVGGRDGRDCEWPMQASI--QHFG---AHVCGGSLIAPQWVLTAAHCFP 127

QY 73 DYPKEDYIVYVGLGRSLNSNTQGENKFEVENLIHKDYSADTLAHNDIALKIRSEGR 132
DB 128 RRALPAEYVRVLGALRGSTSPRTLSVPVRVLLPPDYSED--GARGDLALLQLR----R 181

QY 133 CAQPSRTITQICLPSMYNDPQFGTSCEITGFGKENS'DYLYP-----EQLKMTVVKLI 185
DB 182 PVPLSARVQVCLPVGARPPGTPCRVTGWS-----LRPGVPLPEWPLQGVVPL 235

QY 186 SHRECOQPHYYGVSEVTK-----MLCAADPQWKTSCQDGGPLVCSLQGRMTLTI 238
```

```
DB 236 DSRTCDGLHVAGADVQAERIVLPGLSCAGYPOGHKDACQDGGPLTCLQSGSWVLGV 295
QY 239 VSMGRGKALKDKPGVYTVRVSHFLPWIRS 266
DB 296 VSMGKGKALPNRPGVTVSVATYSPWIOA 323

RESULT 18
ID Q8NF86 PRELIMINARY; PRT; 284 AA.
AC Q8NF86;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Serine protease E0S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Darrow A.L.; Qi J.; Andrade-Gordon P.; Chen C.;
RT "DNA encoding the human serine protease E0S."
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF536382; AAN04055.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE 284 AA; 30110 MW; FDF3F1750D569978 CRC64;

Query Match 28.9%; Score 435.5; DB 4; Length 284;
Best Local Similarity 38.4%; Pred. No. 2.8e-35;
Matches 103; Conservative 35; Mismatches 97; Indels 33; Gaps 8;

QY 13 CGQKTLRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVYVCGGSLISPCWVISAATHCFI 72
DB 28 CGQPRMSSR--IVGGRDGRDCEWPMQASI--QHFG---AHVCGGSLIAPQWVLTAAHCFP 80

QY 73 DYPKEDYIVYVGLGRSLNSNTQGENKFEVENLIHKDYSADTLAHNDIALKIRSEGR 132
DB 81 RRALPAEYVRVLGALRGSTSPRTLSVPVRVLLPPDYSED--GARGDLALLQLR----R 134

QY 133 CAQPSRTITQICLPSMYNDPQFGTSCEITGFGKENS'DYLYP-----EQLKMTVVKLI 185
DB 135 PVPLSARVQVCLPVGARPPGTPCRVTGWS-----LRPGVPLPEWPLQGVVPL 188

QY 186 SHRECOQPHYYGVSEVTK-----MLCAADPQWKTSCQDGGPLVCSLQGRMTLTI 238
DB 189 DSRTCDGLHVAGADVQAERIVLPGLSCAGYPOGHKDACQDGGPLTCLQSGSWVLGV 248

QY 239 VSMGRGKALKDKPGVYTVRVSHFLPWIRS 266
DB 249 VSMGKGKALPNRPGVTVSVATYSPWIOA 276

RESULT 19
ID Q9BK47 PRELIMINARY; PRT; 267 AA.
AC Q9BK47;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Sea Star regeneration-associated protease SRAP.
OC Luidia foliolata.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
```

```
OC Asteroidea; Valvatacea; Paxillosida; Luidiidae; Luidia.
OX NCBI_TaxID=105861;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21100442; PubMed=11179669;
RA Vickery M.C.L.; Vickery M.S., McClintock J.B., Ameller C.D.;
RT "Utilization of a novel deuterostome model for the study of
RT regeneration genetics: Molecular cloning of genes that are
RT differentially expressed during early stages of larval sea star
RT regeneration.";
RL Gene 262:73-80(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF312826; AAK15274.1; --
DR HSSP; P00783; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 267 AA; 28761 MW; 88F61A061921860C CRC64;

Query Match 28.3%; Score 426.5; DB 5; Length 267;
Best Local Similarity 36.2%; Pred. No. 2e-34;
Matches 93; Conservative 42; Mismatches 107; Indels 15; Gaps 8;

QY 13 CGQKTLRPRF-KIIGFEFTTIENQFWFAIYRRHGGSVTVCGGSLISPCWVI 71
DB 18 CGQVINPLNKIVGDEAVPGSWPQVWFRKRYWAGDYQF-CGGTSLISDEWVSAHCF 76
QY 72 IDYPKKEDYIVLGRSLNNTQGMKEFVENLILHKDYSADTLAHHNDAIILKIRSEK 131
DB 77 HNYGINHHTAVVGAHRRDSVDSQTTLVGLKGVFVHESYDSTL--DNDIALIKLSSP-- 132
QY 132 RCAQPSRTIQTICLPSMYNDPQFGSCETITGRG-KENSTDYLPQLKVTYVKLISHREC 190
DB 133 --VSNMNYVNSVCLPTAAPT--GTECVTVGMDQETAVD--DPTLQVVVPIISSEQ 185
QY 191 QPPHYVGSVTTKMLCAADPQWKTSCQDGGPLVC-SLOGRMTLTGIVSGRGCAKD 249
DB 186 NRATVYGGGINNMICAGFKGKGCQDGGPFCVQSASGEYELVGVWSVGYCADAR 245
QY 250 KPGVYTRVSHFLPWIRS 266
DB 246 KPGVYAKVLNYSVINN 262

RESULT 20
Q966V4
ID Q966V4 PRELIMINARY; PRT; 505 AA.
AC Q966V4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Proacrosin.
GN ACR.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyruidae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21326076; PubMed=11313341;
RA Kodama E., Baba T., Yokosawa H., Sawada H.;
RT "cDNA cloning and Functional Analysis of Ascidian Sperm Proacrosin.";
RL J. Biol. Chem. 276:24594-24600(2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AB052635; BAB60718.1; --
DR HSSP; P00761; IANI.
```

```
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 2.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 505 AA; 55002 MW; 79A1A917CE1D9334 CRC64;

Query Match 28.2%; Score 425.5; DB 5; Length 505;
Best Local Similarity 35.5%; Pred. No. 5.8e-34;
Matches 99; Conservative 47; Mismatches 92; Indels 41; Gaps 11;

QY 13 CGQKTLRPRF-----KIIGFEFTTIENQFWFAIYRRHGGSVTVCGGSLISPCWVI 65
DB 21 CG---LRPLOSALITGRIVGGEMAKLGFFPWAALYKH-----VQVCGTITDTTWIL 72
QY 66 SATHCF-----IDYPKKEDYIVYLG-RSLNNTQGMKEFVENLILHKDYSADTLAHH 118
DB 73 SAAHCFDPHMYNLQSIKKEDALIRVADLDKTDDEGEMTFEVKDIILHEQYNRQTF--D 130
QY 119 NDIALIKIRSEKRCQPSSTIQTICLPSMYNDPQFGSCETITGRGKENSTDYL---YPE 175
DB 131 NDIMLIEILGS-----ITYGTVQAPACIPGANDAVDGTCLISGMG--DTQDHVHRWPD 184
QY 176 QLKMTVTKLISHRECOQPHYVGSVTTKMLCAADPQWKTSCQDGGPLVC-----SLQ 230
DB 185 KIQKAQVEVFAQAQCLATY--PESTENMICAGLRTGGIDSCQDGGPLACPTTETAQ 241
QY 231 GRMTLTGIVSGRGCAKOKPGVYTRVSHFLPWIRSHTK 269
DB 242 PTFLOQIVSGRGCAIDGPGFYTVYEVRYSSWIANYTQ 280

RESULT 21
Q9K0D2
ID Q9K0D2 PRELIMINARY; PRT; 517 AA.
AC Q9K0D2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; --
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle; 1.
```

DR	SMART; SMO0191; EGF; 3.
DR	SMART; SMO0130; KR; 1.
DR	SMART; SMO0020; Tryp_Spc; 1.
DR	PROSITE; PS00022; EGF_1; 3.
DR	PROSITE; PS01186; EGF_2; 2.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PS00070; KRINGLE_2; 1.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW	Kringle; Protease; Serine protease.
SQ	SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match	
Best Local Similarity	28.2%; Score 425; DB 11; Length 517;
Matches 101; Conservative	36; Mismatches 102; Indels 22; Gaps 9;

Qy	23 KIIGGEFTTIENQFWFAAY-----RRHRGGSVTYVCGSLSPCWVISATHCFIDYP 75
Db	270 RIYGGFKSTAGKHPQWSLQTSPLTTSMPCG---HFCGGLIHPQWLTAHC-TDIN 324
Qy	76 KKEDIYIVLGSRLSNLTNQEMKEVENILHKDYSDATLAHNDAIALLKIRSKEGRCQA 135
Db	325 TKHLKV-V-LGDQDLKKTESHEQTFRVKILKYQYNRDEIPHNDIAALKLPVGSHCAL 383
Qy	136 PSRTIQICLPSMYNDP-QFGTSCIEITGFGKENSTDLYPEQLAKMTVVKLISHRECCQP 194
Db	384 ESRAVKTCLPS---DFPSPGTECHISGWGVTEGE--GSROLLDAVKLIANPLNSRQ 438
Qy	195 YGSEVTVMKLCADPNWK-TDCQGSGSGLPVCSLOGRWTLTGIVSWGRGCALXKPGV 253
Db	439 LYDHTIDDSMICAGNLQKPSDTCCGDSGGPLTCCKDGTVYVYGIVSWGQECG--KKPGV 496
Qy	254 YTRVSHELPMIRSHTEENGL 274
Db	497 YTQVTKFLNKIKTMHREAGL 517

RESULT 22	
Q99JC8	PRELIMINARY; PRT; 761 AA.
ID Q99JC8	
AC Q99JC8;	01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2003	(TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003	(TrEMBLrel. 23, Last annotation update)
DE Neurotropsin.	
GN NT.	
OS Rattus norvegicus (Rat).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OC NCBI_Taxid=10116;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Hintsch G., Sonderegger P.;	
RT "Cloning and sequencing of the cDNA encoding rat neurotropsin.";	
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.	
DR ENBL; AJ011671; CAC35028.2; -.	
SQ SEQUENCE 761 AA; 81465 MW; 91DC966ED3BACFE1 CRC64;	

Query Match	
Best Local Similarity	28.0%; Score 421.5; DB 11; Length 761;
Matches 98; Conservative	36.6%; Pred. No. 2.5e-33;
	Mismatches 113; Indels 13; Gaps 8;

Zy	7 BELKFQCGKTL-RPRPKIIGGEFTTIENQFWFAAYRRHRGGSVTYVCGSLSPCWVI 65
Db	499 ETLSGCGLRLLHRRQTRIIGSNLSRGAWPWQASLRKSTHGDRLLCGATLLSSCWL 558
Zy	66 SATPCFIDY-PKXEDIVVLGRSLNSNTQGMKFEVENILILHKDYSDATLAHNDAIAL 124
Zy	559 TAAHCTRYYNNRSRYAVRGDVHTLVPEGFQDIGVQVIHNRVRPDSDY--DIALV 616
Zy	125 KIRSKEGRCQAPRTIQTICLPSMYNDPQ-FGTSCITGFGKENSTDLYPEQLKNVTVK 183

```
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=21167393; PubMed=11267681;
RA Kim D.R., Shamin S., Inoue M., Kido H.,
RT "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung."
RL Biochim. Biophys. Acta 1518:204-209 (2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB048797; BAB39742.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.087; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001130; Srcr_receptor.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLA; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00287; SRCR_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 537 AA; 58102 MW; A39FF4E8816DAECF CRC64;

Query Match 27.7%; Score 418; DB 4; Length 537;
Best Local Similarity 36.7%; Pred. No. 3.6e-33;
Matches 97; Conservative 38; Mismatches 105; Indels 24; Gaps 8;

QY 13 CQOKLPRPKFIIGGFTTIENQPMFAAIYRHRGSGVTYVCGSLISPCWVISATHCFI 72
DB 287 CGLRAMTGR--IVGGALASDSKWPQVUSL---HFG--TTHICGGLIDAQWVLTAAHCF 339
QY 73 DVFPKK--EDYIVYGLSRNSNTQGMKEFEVENLIHKDYSADTLAHNDIALLKIRSK 130
DB 340 VTREKVEGCKVYAGTSLNHLQLEPAA--SIAEIIINSYTD--EDDYDIAM----- 388

QY 131 GRCAQP---SRTIOTICLSMYNDPOFGTSCEITGFGKENSDYLYPEOLKMTVVKLISH 187
DB 389 -RLSKLTLTSAIHHPACLPDHGQTFSLNCTCWITGFKTRETDDKTSPLFEVQVNLIDF 447
QY 188 RECOQPHYGVSEVTTKMLCAAPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCAL 247
DB 448 KKCNDVLYVDSYLTTPMPCAGLHGRDSCQDGGPLVCEQNNWYLAGVTSWGTGCGQ 507

QY 248 KKKPGVYTRVSHFLPWIRSHKTEE 271
DB 508 RNKPGVYTKVTEVLPWIPYISKMESE 531

RESULT 25
Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RP [2]
RX SEQUENCE FROM N.A.
RA MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
```

```
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator."
RL J. Biochem. 119:1157-1165 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -.
DR EMBL; S83182; AAB46909.1; -.
DR EMBL; BC031412; AAH31412.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.033; -.
DR Genew; HGNC:4798; HABP2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 27.7%; Score 417; DB 4; Length 560;
Best Local Similarity 36.2%; Pred. No. 4.7e-33;
Matches 104; Conservative 40; Mismatches 113; Indels 30; Gaps 11;

QY 2 PSSPPEELK--FQCGOKTLRPR--FKLIIGGFTTIENQPMFAAIYRHRGSGVT----- 51
DB 288 PTESTKLPDSCGCKTEIAERKIKRIVGFKSTAGKHPQASLQ-----SSLPLTISMP 342

QY 52 --YVCGSLISPCWVISATHCFIDYPKKEDYIVYGLSRNSNTQGMKEFEVENLIHKD 109
DB 343 QGHFCGCGALIHPCWVLTAAHC-TDI-KTRHLKVVILGQDLKKEEFHQSFVEKIFKYSH 400

QY 110 YSADTLAHNDIALLKIRSKEGCAOPSRTIOTICLSMYNDPOF--GTSCETLFGKEN 167
DB 401 YNERDEIPTHNDIALLKUKPDVGHCALESKYVKTVCLP-----DGSFPGSECHISGWYTE 456

QY 168 STDLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWK-TDSCQDGGPLV 226
DB 457 TKG--GSRQLLDAKVKLIANTLNSRLQYLDHMDSDMICAGNLQKPGQDTCQDGGSGPLT 514

QY 227 CSLOGRMTLTGIVSWGRCALKKPGVYTRVSHFLPWIRSHKTEE 273
DB 515 CEKDGTYVYGVISWGLECG--KRPGVYTVQTVFLNWKATIKSESG 559

RESULT 26
Q9BYE2 PRELIMINARY; PRT; 581 AA.
ID Q9BYE2
AC Q9BYE2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Membrane-type mosaic serine protease.
```


DR	InterPro; IPR001254; Ser_protease_Try.
DR	Pfam; PF00057; ldl_recep_c_a; 2.
DR	Blast Local Similarity 27.0%; Score 406.5; DB 11; Length 799;
DR	Matches 99; Conservative 47; Mismatches 89; Indels 35; Gaps 10;
DR	PRINTS; PRO0089; trypsin; 1.
DR	PRINTS; PRO0722; CHYMOTRYPSIN.
DR	PRINTS; PRO0261; LDLRECEPTOR.
DR	PROSITE; PS01180; CUB; 1.
DR	PROSITE; PS01209; LDLRA_1; 1.
DR	PROSITE; PS00068; LDLRA_2; 3.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Hydrolase; Protease; Serine protease.
SW	SEQUENCE 799 AA; 89557 MW; 16315A646A4D5288 CRC64;
Query Match	27.0%; Score 406.5; DB 11; Length 799;
Best Local Similarity	36.7%; Pred. No. 8.4e-32;
Matches	99; Conservative 47; Mismatches 89; Indels 35; Gaps 10;
QY	7 ELKFCQGOKTLRPFKIIGGFTTIENOPWFAAIYRRHRGGSVTVCGSLISPQWVIS 66
DB	550 DEQHCDGGLQLSSR--IVGGTSSGESEGWASLQVGR-----HICGGALIADRWIT 602
QY	67 ATHCFIDYPKED-----YTVYLGSRSLNNTQCEMKFEVENLIHKDYSDATLAHN 119
DB	603 AAHCF-----QEDSMASPKLTWTFVGKNQRNSRPGEVSFKYSRLFLHPVHEED--SHDY 655
QY	120 DIALLKIRSKERCAQP----STIQTICLPSPMYNDPOFGTSCEITGFGKENSTDLYPEQ 176
DB	656 DVALLQL-----DHPVVYSATVPCLPASFHFEPGHCHWITGWGAQREGGPV-SNT 707
QY	177 LKMTVVKLISHRECQPHYGSEVTTKMLCAADPWKTDSQQDSGGPLVC-SLOGRMTL 235
DB	708 LQKVVDVLVPQDLGSEARY--QVSPRMLCAGYRGKKDACQDSDGSLVCPREPSCRWFL 765
QY	236 TGIWSRGRCALKDKPGVTYTRUSHELPWIR 265
DB	766 AGLSVWGLCGRPNFPGVYTRVINWIQ 795
RESULT 30	
Q8IU2	
ID	Q8IU2 PRELIMINARY; PRT; 802 AA.
AC	Q8IU2,
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
OS	Matriptase-2.
DE	Os Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Ummalialia; Eutheria; Primates; Catarhini; Hominidae; Homo.
OX	NCBI_Taxid=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Liver;
RX	MEDLINE=22241917; PubMed=12149247;
RT	Velasco G., Cal S., Quesada V., Sanchez L.M., Lopez-Otin C.;
RT	"Matriptase-2, a membrane-bound mosaic serine proteinase predominantly
RT	expressed in human liver and showing degrading activity against
RT	extracellular matrix proteins.";
RL	J. Biol. Chem. 277:37637-37646(2002).
RL	EMBL; AJ319876; CAD85953.1; .
SQ	SEQUENCE 802 AA; 88901 MW; C30D37BEC4FD22E CRC64;
Query Match	26.8%; Score 404.5; DB 4; Length 802;
Best Local Similarity	36.8%; Pred. No. 1.3e-31;
Matches	99; Conservative 48; Mismatches 89; Indels 33; Gaps 10;
QY	7 ELKFCQGOKTLRPFKIIGGFTTIENOPWFAAIYRRHRGGSVTVCGSLISPQWVIS 66
DB	553 DEEHCDGGLQ--GPSSRIYGGAVSGEWASLQVRG-----HICGGALIADRWIT 605
QY	67 ATHCFIDYPKED-----YTVYLGSRSLNNTQCEMKFEVENLIHKDYSDATLAHN 119

Db 606 AHCF-----QEDSMASVTWTVFLGVWQNSRWPGVSKVSRLLLPYHEED--SHDY 658
Qy 120 DIALKIRSKGRCAOPSRITQITICLPMSWYNDPQFOTSCIEITFG--KENSTDYLYPEOL 177
Db 659 DVALLQLDHPVVR-----SAAVRPVCLPARSHFFEPGLHCWITGAGALREGGP---ISNAL 711
Qy 178 KMTVVKLISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDGGGLVC-SIQGRMTLT 236
Db 712 QKVDVQLIPQDLCSEVRY--QVTPRLCAGYRKGXKDACQDGGGLVCALSGRWFLA 769
Qy 237 GIVSWGRCALKDKPGVYTRVSHFLPWIR 265
Db 770 GLVSWGIGCRPNYFGVYTRITGVISWIQ 798

Search completed: January 8, 2004, 13:02:07
Job time : 44 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2004, 13:17:56 ; Search time 41 Seconds

(without alignments)
1068.501 Million cell updates/sec

Title: US-09-880-503-5

Perfect score: 1508
Sequence: 1 KPSSPPBLKFOCQKTLRP.....VSHFLPWIRSHTKENGLAL 276

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107663 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 74

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 100%
Listing first 250 summaries

Database: A Geneseq 19Jun03:*

```
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1508	100.0	276	23	AAE16546	Human LMW-urokinas
2	1508	100.0	323	23	AAE16548	Human uPA deltakri
3	1508	100.0	365	16	AAE68854	Delta 1-46 urokina
4	1508	100.0	386	16	AAE68266	Bifunctional uroki
5	1508	100.0	390	16	AAE66245	Bifunctional uroki
6	1508	100.0	390	16	AAE66247	Bifunctional uroki
7	1508	100.0	392	16	AAE66254	Bifunctional uroki
8	1508	100.0	392	16	AAE66255	Bifunctional uroki
9	1508	100.0	392	16	AAE66256	Bifunctional uroki

ALIGNMENTS

RESULT 1
AAE16546
ID AAE16546 standard; Protein; 276 AA.

10	1508	100.0	392	16	AAE66257	Bifunctional uroki
11	1508	100.0	392	16	AAE66258	Bifunctional uroki
12	1508	100.0	392	16	AAE66259	Bifunctional uroki
13	1508	100.0	392	16	AAE66260	Bifunctional uroki
14	1508	100.0	392	16	AAE66261	Bifunctional uroki
15	1508	100.0	392	16	AAE66263	Bifunctional uroki
16	1508	100.0	392	16	AAE66264	Bifunctional uroki
17	1508	100.0	393	16	AAE66244	Bifunctional uroki
18	1508	100.0	393	16	AAE66249	Bifunctional uroki
19	1508	100.0	393	16	AAE66250	Bifunctional uroki
20	1508	100.0	393	16	AAE66251	Bifunctional uroki
21	1508	100.0	393	16	AAE66252	Bifunctional uroki
22	1508	100.0	393	16	AAE66253	Bifunctional uroki
23	1508	100.0	393	17	AAE99885	M36: fibrinolytic
24	1508	100.0	393	17	AAE99896	Chimeric protein M
25	1508	100.0	393	17	AAE99897	Chimeric protein M
26	1508	100.0	395	16	AAE66262	Bifunctional uroki
27	1508	100.0	395	16	AAE66265	Bifunctional uroki
28	1508	100.0	396	16	AAE66246	Bifunctional uroki
29	1508	100.0	397	16	AAE66248	Bifunctional uroki
30	1508	100.0	411	6	AAE50871	Sequence encoded b
31	1508	100.0	411	10	AAE93589	Amino acid sequenc
32	1508	100.0	411	10	AAE96146	Sequence encoded b
33	1508	100.0	411	10	AAE91684	Sequence of urokin
34	1508	100.0	411	11	AAE6244	Urokinase precursor
35	1508	100.0	411	11	AAE07902	Human pro-urokinas
36	1508	100.0	411	11	AAE07903	Human pro-urokinas
37	1508	100.0	411	11	AAE07904	Human pro-urokinas
38	1508	100.0	411	14	AAE34584	Mutant human prou
39	1508	100.0	411	16	AAE62991	Pro-urokinase. Ho
40	1508	100.0	411	20	AAE39343	Human pro-urokinas
41	1508	100.0	411	20	AAE42284	Human pro-urokinas
42	1508	100.0	411	21	AAE92836	Urokinase plasmin
43	1508	100.0	411	22	AAE20489	Human pro-urokinas
44	1508	100.0	411	22	AAE74797	Prourokinase prote
45	1508	100.0	421	23	AAE16544	Human urokinase-ty
46	1508	100.0	421	12	AAE10172	Human prourokinase
47	1508	100.0	431	6	AAE50114	Sequence encoded b
48	1508	100.0	431	7	AAE60674	Modified human pro
49	1508	100.0	431	7	AAE60783	Human urokinase.
50	1508	100.0	431	8	AAE70258	Sequence of A- and
51	1508	100.0	431	8	AAE70250	Deduced AA sequenc
52	1508	100.0	431	9	AAE80430	Pro-urokinase with
53	1508	100.0	431	9	AAE81204	Natural human prou
54	1508	100.0	431	10	AAE92119	Human pro-urokinas
55	1508	100.0	431	11	AAE04253	Human pro-urokinas
56	1508	100.0	431	11	AAE07112	Human pro-urokinas
57	1508	100.0	431	15	AAE47903	Pro-urokinase deri
58	1508	100.0	431	15	AAE63141	Full length human
59	1508	100.0	431	21	AAE99591	Human plasminogen
60	1508	100.0	431	21	AAE50869	Human urokinase pr
61	1508	100.0	431	22	AAE84605	Amino acid sequenc
62	1508	100.0	431	23	AAE79460	Human plasminogen
63	1508	100.0	431	23	AAE99228	Human plasminogen
64	1508	100.0	431	23	AAE99232	Human plasminogen
65	1508	100.0	431	23	AAE99236	Human plasminogen
66	1508	100.0	431	23	AAE99240	Human plasminogen
67	1508	100.0	431	23	AAE17128	Human uPA protein.
68	1508	100.0	431	24	ABU56547	Lung cancer-associ
69	1508	100.0	431	24	ABU56708	Lung cancer-associ
70	1508	100.0	431	24	ABU11076	Human urokinase p1
71	1508	100.0	432	17	AAE99888	M43: fibrinolytic
72	1508	100.0	434	13	AAE20537	Amidated deriv. of
73	1508	100.0	434	13	AAE20538	Amidated deriv. of
74	1508	100.0	436	13	AAE20536	Amidated deriv. of

XX AAE16546;
AC
XX
DT 09-APR-2002 (first entry)
XX
XX Human LMW-urokinase-type plasminogen activator (uPA) protein.
XX
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW male impotence; adult respiratory distress syndrome; LMW-uPA.
XX
XX Homo sapiens.
XX
XX WO200197752-A2.
XX
XX 27-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-US18976.
XX
XX 20-JUN-2000; 2000US-212874P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Cines DB, Higazi AA;
XX
XX WPI; 2002-122240/16.
XX
XX N-PSDB; AAD27079.
XX
XX Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
XX Claim 5; Fig 1E; 117pp; English.
XX
XX The invention relates to a composition comprising one or more domains of
XX urokinase-type plasminogen activator (uPA). The composition is used to
XX modulate the contractility and angiogenic activity of a mammalian muscle,
XX endothelial cell or tissue. The composition is used for treating stroke,
XX hypotension, hypertension, atherosclerosis, heart attack, microvascular
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic
XX disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
XX diabetic retinopathy, wound healing, clotting disorder, uterine
XX contraction disorder, male impotence, respiratory disease or condition
XX such as asthma, adult respiratory distress syndrome, primary pulmonary
XX hypertension, microvascular thrombotic occlusion, and a disorder
XX associated with chronic intrapulmonary fibrin formation. The present
XX sequence is human LMW-urokinase-type plasminogen activator (uPA) protein.
XX
SQ Sequence 276 AA;
Query Match 100.0%; Score 1508; DB 23; Length 276;
Best Local Similarity 100.0%; Pred. No. 1.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFCQCKTLPKFIIGGFTTIEQNPFAAIYRHRGGSVTVYCGSLIS 60
DB 1 KPSSPPEELKFCQCKTLPKFIIGGFTTIEQNPFAAIYRHRGGSVTVYCGSLIS 60
QY 61 PCWVISATHCFIDYPKKEDYVILGRSLNNTQGMKFVENILHKDYSAADTLAHND 120
DB 61 PCWVISATHCFIDYPKKEDYVILGRSLNNTQGMKFVENILHKDYSAADTLAHND 120
QY 121 IALLKIRSEGRCAQPSRTIOTICLPSPYNDPQFGTSCEITGPKENSTDYLYPEQLKMT 180
DB 121 IALLKIRSEGRCAQPSRTIOTICLPSPYNDPQFGTSCEITGPKENSTDYLYPEQLKMT 180
QY 181 VVKLISHRECCQPHYVYGSVTTKMLCAADPQWKTDSCQDSGGLVCSLQGRMTLTGIVS 240
DB 181 VVKLISHRECCQPHYVYGSVTTKMLCAADPQWKTDSCQDSGGLVCSLQGRMTLTGIVS 240

QY 241 WGRGCALCKDKPGVTVRVSHFLPWIRSHTKENGLAL 276
DB 241 WGRGCALCKDKPGVTVRVSHFLPWIRSHTKENGLAL 276
RESULT 2
AAE16548
ID AAE16548 standard; Protein; 323 AA.
XX
XX AAE16548;
AC
XX 09-APR-2002 (first entry)
DT
XX
XX Human uPA deltakringle-scuPA and deltakringle-tcuPA mutant.
DE
XX
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW male impotence; adult respiratory distress syndrome; tcuPA; scuPA;
KW two chain urokinase; single chain urokinase; mutant; mutein.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX WO200197752-A2.
XX
XX 27-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-US18976.
XX
XX 20-JUN-2000; 2000US-212874P.
PR
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Cines DB, Higazi AA;
XX
XX WPI; 2002-122240/16.
XX
XX N-PSDB; AAD27081.
XX
XX Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
XX Claim 23; Fig 1G; 117pp; English.
XX
XX The invention relates to a composition comprising one or more domains of
XX urokinase-type plasminogen activator (uPA). The composition is used to
XX modulate the contractility and angiogenic activity of a mammalian muscle,
XX endothelial cell or tissue. The composition is used for treating stroke,
XX hypotension, hypertension, atherosclerosis, heart attack, microvascular
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic
XX disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
XX diabetic retinopathy, wound healing, clotting disorder, uterine
XX contraction disorder, male impotence, respiratory disease or condition
XX such as asthma, adult respiratory distress syndrome, primary pulmonary
XX hypertension, microvascular thrombotic occlusion, and a disorder
XX associated with chronic intrapulmonary fibrin formation. The present
XX sequence is human urokinase-type plasminogen activator (uPA) deltakringle
XX single chain urokinase (scuPA) and deltakringle-two chain urokinase
XX (tcuPA) deletion mutant.
XX
SQ Sequence 323 AA;
Query Match 100.0%; Score 1508; DB 23; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.7e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFCQCKTLPKFIIGGFTTIEQNPFAAIYRHRGGSVTVYCGSLIS 60

Db 48 KPSSPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFAPFAIYRRHGGSVTVVCGGSLIS 107

Qy 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120

Db 108 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 167

Qy 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDYLYPEQLKMT 180

Db 168 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDYLYPEQLKMT 227

Qy 181 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCSQGSGGPLVCSLQGRMTLTGIVS 240

Db 228 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCSQGSGGPLVCSLQGRMTLTGIVS 287

Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 288 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 323

RESULT 3

ID AAR68854 standard; protein; 365 AA.

AC AAR68854;

XX

DT 25-MAR-2003 (updated)

DT 22-NOV-1995 (first entry)

XX

DE Delta 1-46 urokinase.

XX

XX Human; des-epidermal growth factor homologous plasminogen activator;

KW uPA; liver membrane; reduced affinity; BGF homologous; thrombosis;

KW thrombolytic; increased half-life; urokinase.

XX

OS Homo sapiens (engineered).

XX

PH Key Location/Qualifiers

FT Misc-difference 1

FT /note= "amino acids 1-46 of wild-type urokinase

FT have been deleted"

XX

PN US376547-A.

XX

PD 27-DEC-1994.

XX

PF 29-JAN-1988; 88US-0150267.

XX

PR 30-JAN-1987; 87US-0008795.

PR 29-JAN-1988; 88US-0150267.

XX

PA (AMHP) AMERICAN HOME PROD CORP.

XX

PI Hung PP, Kalyan NK, Lee SL;

XX

DR WPI; 1995-043464/06.

XX

XX New modified plasminogen activator cpds. - having regions removed

PT to reduce affinity for liver membranes and increase circulation

PT half-life.

XX

PS Claim 1; Page ?; 25pp; English.

XX

CC Amino acid residues 1-46 contain the EGF region of human urokinase.

CC Deletion of this region results in a plasminogen activator with

CC reduced affinity for liver cell membranes; the mutant protein is

CC not cleared from the circulation as rapidly as is wild-type tPA.

CC The specification only gives the sequence around the deletion and

CC not the full-length sequence of "delta 1-46 urokinase"; the

CC sequence in AAR68854 has been obtained by amending a previously

CC disclosed wild-type human urokinase sequence (from W09501427)

CC according to the description given in Example 3.

CC (Updated on 25-MAR-2003 to correct PF field.)

XX

SQ Sequence 365 AA;

Query Match 100.0%; Score 1508; DB 16; Length 365;

Best Local Similarity 100.0%; Pred. No. 1.9e-125; Indels 0; Gaps 0;

Matches 276; Conservative 0; Mismatch 0;

Qy 1 KPSSPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFAPFAIYRRHGGSVTVVCGGSLIS 60

Db 90 KPSSPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFAPFAIYRRHGGSVTVVCGGSLIS 149

Qy 61 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120

Db 150 PCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 209

Qy 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDYLYPEQLKMT 180

Db 210 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDYLYPEQLKMT 269

Qy 181 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCSQGSGGPLVCSLQGRMTLTGIVS 240

Db 270 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCSQGSGGPLVCSLQGRMTLTGIVS 329

Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 4

ID AAR66266 standard; protein; 386 AA.

AC AAR66266;

XX

DT 25-MAR-2003 (updated)

DT 17-AUG-1995 (first entry)

XX

DE Bifunctional urokinase variant M33.

XX

KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

KW urokinase; variant; mutein.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Region 1..365

FT /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4..85

FT Disulfide-bond 25..67

FT Disulfide-bond 56..80

FT Disulfide-bond 102..233

FT Disulfide-bond 143..159

FT Disulfide-bond 151..222

FT Disulfide-bond 247..316

FT Disulfide-bond 279..295

FT Disulfide-bond 306..334

FT Disulfide-bond 365..366

FT Region /label= X1

FT /note= "peptide bond"

FT 366..386

FT Region /label= Y1

XX

PN DE4323754-C1.

XX

XX 01-DEC-1994.

XX

PF 15-JUL-1993; 93DE-4323754.

XX

PR 15-JUL-1993; 93DE-4323754.

XX

PA (CHEF) GRUENENTHAL GMBH.

XX

PI Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;

DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M14.
 DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 XX urokinase; variant; mutein.
 KW
 XX
 OS Synthetic.
 XX
 XX
 PH Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..390
 FT /label= Y1
 XX
 XX DE4323754-C1.
 XX
 XX 01-DEC-1994.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wendt S;
 XX
 XX WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derive and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 XX Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 390 AA;
 SQ
 Query Match 100.0%; Score 1508; DB 16; Length 390;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPEELKFCQGKTLPRFKIIGGETTIENOPWFAAIYRRRGSGSVTVVCGSLIS 60
 DB 90 KPSSPEELKFCQGKTLPRFKIIGGETTIENOPWFAAIYRRRGSGSVTVVCGSLIS 149
 QY 61 PCWVSATHCFTDPKEDYIYVILGRSLNSNTQGMKEFVENILHKDYSADTLAHND 120
 DB 150 PCWVSATHCFTDPKEDYIYVILGRSLNSNTQGMKEFVENILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAOPSRITQITCLPSMYNDPQFGTSCEITGFGKENSTDYLYPQLKMT 180

DB 210 IALLKIRSEGRCAOPSRITQITCLPSMYNDPQFGTSCEITGFGKENSTDYLYPQLKMT 269
 QY 181 VVKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQGDSCGGLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQGDSCGGLVCSLQGRMTLTGIVS 329
 QY 241 WGRGCALCKDPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGCALCKDPGVYTRVSHFLPWIRSHTKENGLAL 365
 RESULT 7
 AAR66254
 ID AAR66254 standard; protein; 392 AA.
 XX
 AC AAR66254;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M21.
 XX
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 XX
 FT Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 FT /label= Y1
 XX
 XX DE4323754-C1.
 XX
 XX 01-DEC-1994.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wendt S;
 XX
 XX WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derive and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 XX Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 390 AA;
 SQ
 Query Match 100.0%; Score 1508; DB 16; Length 390;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPEELKFCQGKTLPRFKIIGGETTIENOPWFAAIYRRRGSGSVTVVCGSLIS 60
 DB 90 KPSSPEELKFCQGKTLPRFKIIGGETTIENOPWFAAIYRRRGSGSVTVVCGSLIS 149
 QY 61 PCWVSATHCFTDPKEDYIYVILGRSLNSNTQGMKEFVENILHKDYSADTLAHND 120
 DB 150 PCWVSATHCFTDPKEDYIYVILGRSLNSNTQGMKEFVENILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAOPSRITQITCLPSMYNDPQFGTSCEITGFGKENSTDYLYPQLKMT 180

CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX
SQ Sequence 392 AA;

Query Match 100.0%; Score 1508; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.1e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60
DB 90 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 149

QY 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209

QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 269

QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 365

RESULT 8
AAR66255
ID AAR66255 standard; protein; 392 AA.
XX
XX
AC AAR66255;
XX
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M22.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT 372..392
FT Region /label= Y1
XX
XX
PN DE4323754-Cl.
XX
XX
PD 01-DEC-1994.
XX
XX
PF 15-JUL-1993; 93DE-4323754.
XX
XX
PR 15-JUL-1993; 93DE-4323754.
XX

(CHEF) GRUENENTHAL GMBH.
XX
XX
PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
XX
WPI; 1995-015191/03.
XX
DR
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs, which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX
SQ Sequence 392 AA;

Query Match 100.0%; Score 1508; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.1e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60
DB 90 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 149

QY 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209

QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 269

QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 365

RESULT 9
AAR66256
ID AAR66256 standard; protein; 392 AA.
XX
XX
AC AAR66256;
XX
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M23.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80

FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 FT /label= Y1
 XX
 PN DE4323754-Cl.
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wendt S;
 XX
 XX WPI; 1995-015191/03.
 XX
 DR New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 392 AA;
 Query Match 100.0%; Score 1508; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVCGSLIS 60
 DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVCGSLIS 149
 QY 61 PCWVISATHCFIDYPKKEDYIVYLSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWVISATHCFIDYPKKEDYIVYLSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRKEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGKENSVDYLYPEQLKMT 180
 DB 210 IALLKIRKEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGKENSVDYLYPEQLKMT 269
 QY 181 VVKLISHRECQPHYGYSEVITKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECQPHYGYSEVITKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365
 RESULT 10
 AAR66257
 ID AAR66257 standard; protein; 392 AA.
 XX

AC AAR66257;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M24.
 XX
 KW fibrinolytic; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FH Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 FT /label= X1
 FT Region 372..392
 FT /label= Y1
 XX
 PN DE4323754-Cl.
 XX
 PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wendt S;
 XX
 XX WPI; 1995-015191/03.
 XX
 DR New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 392 AA;
 Query Match 100.0%; Score 1508; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2.1e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVCGSLIS 60
 DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVCGSLIS 149
 QY 61 PCWVISATHCFIDYPKKEDYIVYLSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWVISATHCFIDYPKKEDYIVYLSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRKEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGKENSVDYLYPEQLKMT 180
 DB 210 IALLKIRKEGRCAQPSRTIQTICLPSMYNDPQFTSCITGFGKENSVDYLYPEQLKMT 269
 QY 181 VVKLISHRECQPHYGYSEVITKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECQPHYGYSEVITKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365
 RESULT 10
 AAR66257
 ID AAR66257 standard; protein; 392 AA.
 XX

Db 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 209
Qy 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
Db 210 IALLKIRSKGRCAPSRITQICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 269
Qy 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSQGDGSGPLVCSLQGRMTLTGIVS 240
Db 270 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSQGDGSGPLVCSLQGRMTLTGIVS 329
Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365
RESULT 11
AAR66258
ID AAR66258 standard; protein; 392 AA.
XX
AC AAR66258;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M25.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT Disulfide-bond 4..85 /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX
XX DB4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF) GRUENTHAL GMBH.
XX
XX Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved

CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 392 AA;
Query Match 100.0%; Score 1508; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.le-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 KPSSPPBELKFCQCKTLRPRFKIIGGEFTTIENQPFALYRHRGGSVTVCGSLIS 60
Db 90 KPSSPPBELKFCQCKTLRPRFKIIGGEFTTIENQPFALYRHRGGSVTVCGSLIS 149
Qy 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
Db 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 209
Qy 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
Db 210 IALLKIRSKGRCAPSRITQICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 269
Qy 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSQGDGSGPLVCSLQGRMTLTGIVS 240
Db 270 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSQGDGSGPLVCSLQGRMTLTGIVS 329
Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365
RESULT 12
AAR66259
ID AAR66259 standard; protein; 392 AA.
XX
AC AAR66259;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M26.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT Disulfide-bond 4..85 /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX


```

XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF ) GRUENENTHAL GMBH.
XX PA Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX DR WPI; 1995-015191/03.
XX XX New bifunctional urokinase derivs and related plasmids - with
XX FT improved fibrinolytic and thrombin inhibiting activities, for
XX FT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX XX
XX PS Example 1; Page 11 and Fig 1; 34pp; German.
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX XX
XX SQ Sequence 392 AA;
XX
Query Match 100.0%; Score 1508; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.1e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFCQCKQTLRPRFKIIGGFTTIENQPFAPAIYRRHRGGSVTVVCGSLIS 60
DB 90 KPSSPPEELKFCQCKQTLRPRFKIIGGFTTIENQPFAPAIYRRHRGGSVTVVCGSLIS 149
QY 61 PCWVISATHCFIDYPKKEDYVILGRSLNSQTQGMKFEVENLIHKDYSADTLAHND 120
DB 150 PCWVISATHCFIDYPKKEDYVILGRSLNSQTQGMKFEVENLIHKDYSADTLAHND 209
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
XX
RESULT 13
ID AAR66260 standard; protein; 392 AA.
XX AC AAR66260;
XX AC AAR66260;
XX DT 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M27.
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; muten.
XX CS Synthetic.
XX FH Key Location/Qualifiers
XX FT Region 1..365
XX FT /label= M4
XX FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

```

```

FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 386..371
FT /label= X1
FT Region 372..392
FT /label= Y1
XX
XX DB4323754-CI.
XX
XX PD 01-DEC-1994.
XX
XX PF 15-JUL-1993; 93DE-4323754.
XX
XX PR 15-JUL-1993; 93DE-4323754.
XX
XX PA (CHEF ) GRUENENTHAL GMBH.
XX
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wnendt S;
XX
XX DR WPI; 1995-015191/03.
XX
XX XX New bifunctional urokinase derivs and related plasmids - with
XX FT improved fibrinolytic and thrombin inhibiting activities, for
XX FT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX XX
XX PS Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 392 AA;
XX
Query Match 100.0%; Score 1508; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.1e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFCQCKQTLRPRFKIIGGFTTIENQPFAPAIYRRHRGGSVTVVCGSLIS 60
DB 90 KPSSPPEELKFCQCKQTLRPRFKIIGGFTTIENQPFAPAIYRRHRGGSVTVVCGSLIS 149
QY 61 PCWVISATHCFIDYPKKEDYVILGRSLNSQTQGMKFEVENLIHKDYSADTLAHND 120
DB 150 PCWVISATHCFIDYPKKEDYVILGRSLNSQTQGMKFEVENLIHKDYSADTLAHND 209
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
XX
XX WGRGALKDKPGVTVRVSHFLPWIRSHTKKEENGLAL 276
XX WGRGALKDKPGVTVRVSHFLPWIRSHTKKEENGLAL 365

```

RESULT 14

```
AAR66261
ID AAR66261 standard; protein; 392 AA.
AC AAR66261;
XX
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M28.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..392
FT /label= Y1
XX
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 1508; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 2..e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KPSSPPELKFQCGOKTLRPFKLIIGGFTTIENQPFALYRRHGGSVTYVCGSLIS 60
XX
XX 90 KPSSPPELKFQCGOKTLRPFKLIIGGFTTIENQPFALYRRHGGSVTYVCGSLIS 149
XX
```

```
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEEVENLILHKDYSADTLAHHND 120
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEEVENLILHKDYSADTLAHHND 209
QY 121 IALLKIRSKEGRCAPQSRITQITICLPMSYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKEGRCAPQSRITQITICLPMSYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRCQQPHYGGSEVTTKMLCAADPQWKTDSCQSDSGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRCQQPHYGGSEVTTKMLCAADPQWKTDSCQSDSGGPLVCSLQGRMTLTGIVS 329
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 15
AAR66263
ID AAR66263 standard; protein; 392 AA.
XX
AC AAR66263;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M30.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..365
XX /label= M4
XX Disulfide-bond 4..85
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX Disulfide-bond 102..233
XX Disulfide-bond 143..159
XX Disulfide-bond 151..222
XX Disulfide-bond 247..316
XX Disulfide-bond 279..295
XX Disulfide-bond 306..334
XX Disulfide-bond 366..371
XX Region /label= X1
XX Region 372..392
XX /label= Y1
XX
XX DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 1508; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 2..e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KPSSPPELKFQCGOKTLRPFKLIIGGFTTIENQPFALYRRHGGSVTYVCGSLIS 60
XX
XX 90 KPSSPPELKFQCGOKTLRPFKLIIGGFTTIENQPFALYRRHGGSVTYVCGSLIS 149
XX
```

CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 392 AA;
 XX
 Query Match 100.0%; Score 1508; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2.le-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPFPAALYRRHGGSVTVVCGSLIS 60
 DB 90 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPFPAALYRRHGGSVTVVCGSLIS 149
 QY 61 PCWVISATHCFIDYPKEDYIVVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWVISATHCFIDYPKEDYIVVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFSGTSCBITGFKENSTDYLYPEQLKMT 180
 DB 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFSGTSCBITGFKENSTDYLYPEQLKMT 269
 QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 RESULT 16
 AAR66264
 ID AAR66264 standard; protein; 392 AA.
 AC AAR66264;
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 DE Bifunctional urokinase variant M31.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Region 1..365 /label= M4
 FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
 FT 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Disulfide-bond 366..371
 FT Region /label= X1
 FT 372..392
 FT Region /label= Y1
 XX
 DE4323754-Cl.

PD 01-DEC-1994.
 XX
 PF 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 PA (CHEF) GRUENENTHAL GMBH.
 XX
 PI Heinkel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX
 DR WPI; 1995-015191/03.
 XX
 PT New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 392 AA;
 XX
 Query Match 100.0%; Score 1508; DB 16; Length 392;
 Best Local Similarity 100.0%; Pred. No. 2.le-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPFPAALYRRHGGSVTVVCGSLIS 60
 DB 90 KPSSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQPFPAALYRRHGGSVTVVCGSLIS 149
 QY 61 PCWVISATHCFIDYPKEDYIVVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWVISATHCFIDYPKEDYIVVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFSGTSCBITGFKENSTDYLYPEQLKMT 180
 DB 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFSGTSCBITGFKENSTDYLYPEQLKMT 269
 QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
 RESULT 17
 AAR66244
 ID AAR66244 standard; protein; 393 AA.
 AC AAR66244;
 XX
 DT 25-MAR-2003 (updated)
 DT 22-AUG-1995 (first entry)
 DE Bifunctional urokinase variant M11.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers

QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYVCGSLIS 60
 DB 90 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYVCGSLIS 149
 QY 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFTSCETIYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 180
 DB 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFTSCETIYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 269
 QY 181 VVKLIHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLIHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGALKDKPGVTVRVSHFLPWIRSHTEKENGIAL 276
 DB 330 WGRGALKDKPGVTVRVSHFLPWIRSHTEKENGIAL 365

RESULT 19
 AAR66250
 ID AAR66250 standard; protein; 393 AA.
 AC AAR66250;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M17.
 XX
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.

Key Location/Qualifiers
 Region 1..365 /label= M4
 Disulfide-bond 4..85
 Disulfide-bond 25..67
 Disulfide-bond 56..80
 Disulfide-bond 102..233
 Disulfide-bond 143..159
 Disulfide-bond 151..222
 Disulfide-bond 247..316
 Disulfide-bond 279..295
 Disulfide-bond 306..334
 Region 366..372 /label= X1
 Region 373..393 /label= Y1

DE4323754-Cl.
 XX
 XX 01-DEC-1994.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENENTHAL GMBH.
 XX
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX
 DR WPI; 1995-015191/03.
 XX
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc

XX Example 1; Page 10 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX

Sequence 393 AA;
 Query Match 100.0%; Score 1508; DB 16; Length 393;
 Best Local Similarity 100.0%; Pred. No. 2..1e-125; Indels 0; Gaps 0;
 Matches 276; Conservative 0; Mismatches 0;
 QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYVCGSLIS 60
 DB 90 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPMFAAIYRRHRGGSVTVYVCGSLIS 149
 QY 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209
 QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFTSCETIYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 180
 DB 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFTSCETIYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 269
 QY 181 VVKLIHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGPLVCSLQGRMTLTGIVS 240
 DB 270 VVKLIHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGPLVCSLQGRMTLTGIVS 329
 QY 241 WGRGALKDKPGVTVRVSHFLPWIRSHTEKENGIAL 276
 DB 330 WGRGALKDKPGVTVRVSHFLPWIRSHTEKENGIAL 365

RESULT 20
 AAR66251
 ID AAR66251 standard; protein; 393 AA.
 AC AAR66251;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX
 DE Bifunctional urokinase variant M18.
 XX
 KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW urokinase; variant; mutein.
 XX
 OS Synthetic.
 Key Location/Qualifiers
 Region 1..365 /label= M4
 Disulfide-bond 4..85
 Disulfide-bond 25..67
 Disulfide-bond 56..80
 Disulfide-bond 102..233
 Disulfide-bond 143..159
 Disulfide-bond 151..222
 Disulfide-bond 247..316
 Disulfide-bond 279..295
 Disulfide-bond 306..334
 Region 366..372 /label= X1
 Region 373..393 /label= Y1

```

XX PN DE4323754-C1.
XX PD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX PA (CHEF ) GRUENENTHAL GMBH.
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wndt S;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 393 AA;
XX
XX Query Match 100.0%; Score 1508; DB 16; Length 393;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGFTTINQWPFAAIYRRHGGSVTVYCGSLIS 60
DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGFTTINQWPFAAIYRRHGGSVTVYCGSLIS 149
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHDKYSADTLAHND 120
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHDKYSADTLAHND 209
QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
XX
RESULT 21
ID AAR66252
XX AAR66252 standard; protein; 393 AA.
XX AC AAR66252;
XX
XX 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX DE Bifunctional urokinase variant M19.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.

```

```

OS Synthetic.
XX Key Location/Qualifiers
XX Region 1..365 /label= M4
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
XX Disulfide-bond 4..85
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX Disulfide-bond 102..233
XX Disulfide-bond 143..159
XX Disulfide-bond 151..222
XX Disulfide-bond 247..316
XX Disulfide-bond 279..295
XX Disulfide-bond 306..334
XX Disulfide-bond 366..372
XX Region /label= X1
XX Region 373..393 /label= Y1
XX DB4323754-C1.
XX
XX DD 01-DEC-1994.
XX PF 15-JUL-1993; 93DE-4323754.
XX PR 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wndt S;
XX WPI; 1995-015191/03.
XX
XX DR New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 393 AA;
XX
XX Query Match 100.0%; Score 1508; DB 16; Length 393;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGFTTINQWPFAAIYRRHGGSVTVYCGSLIS 60
DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGFTTINQWPFAAIYRRHGGSVTVYCGSLIS 149
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHDKYSADTLAHND 120
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHDKYSADTLAHND 209
QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

```

DE Human pro-urokinase variant.
 XX Thrombin; fibrin; bleeding; pHR24.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Domain 10..42
 FT /label= Epidermal growth factor (EGF) domain
 FT Active-site 24..26
 FT /label= Modified site
 FT Region 10..19
 FT /label= First loop
 FT Region 20..31
 FT /label= Second loop
 FT Region 33..42
 FT /label= Third loop
 XX EP398362-A.
 PN 22-NOV-1990.
 XX 18-MAY-1990; 90EP-0109473.
 XX 18-MAY-1989; 89JP-0126434.
 XX (GRC) GREEN CROSS CORP.
 XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M;
 PI Kawabe H, Arimura H;
 XX WPI; 1990-350147/47.
 DR N-PSDB; AAQ06134.
 XX Human pro-urokinase variant - produced by recombinant methods,
 PT showing increased half life in blood and high affinity for fibrin.
 XX Disclosure; Fig 1; 27pp; English.
 XX Modified pro-urokinase has a longer half-life in blood, and dissolves
 CC thrombin without causing the spontaneous bleeding associated with
 CC urokinase. The modification puts an epidermal growth factor domain
 CC into the protein, containing the sequence Asn-X-Ser or Asn-X-Thr
 CC where X is any residue.
 CC Plasmid pHR24 is disclosed as containing the modified sequence.
 XX Sequence 411 AA;
 Query Match 100.0%; Score 1508; DB 11; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVYCGSLIS 60
 DB 136 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVYCGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCETIGFGKENSVDLYPEQLKMT 180
 DB 256 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCETIGFGKENSVDLYPEQLKMT 315
 QY 181 VVKLISHRECOQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 316 VVKLISHRECOQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGCALKDKEGVYTVRVSHFELPIRSHTKENGLAL 276
 DB 376 WGRGCALKDKEGVYTVRVSHFELPIRSHTKENGLAL 411

RESULT 37
 AAR07904
 ID AAR07904 standard; protein; 411 AA.
 XX
 XX AAR07904;
 XX 21-FEB-1991 (first entry)
 DT Human pro-urokinase variant.
 XX Thrombin; fibrin; bleeding; pHR27.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Domain 10..42
 FT /label= Epidermal growth factor (EGF) domain
 FT Active-site 27..29
 FT /label= Modified site
 FT Region 10..19
 FT /label= First loop
 FT Region 20..31
 FT /label= Second loop
 FT Region 33..42
 FT /label= Third loop
 XX EP398362-A.
 PN 22-NOV-1990.
 XX 18-MAY-1990; 90EP-0109473.
 XX 18-MAY-1989; 89JP-0126434.
 XX (GRC) GREEN CROSS CORP.
 XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M;
 PI Kawabe H, Arimura H;
 XX WPI; 1990-350147/47.
 DR N-PSDB; AAQ06135.
 XX Human pro-urokinase variant - produced by recombinant methods,
 PT showing increased half life in blood and high affinity for fibrin.
 XX Disclosure; Fig 1; 27pp; English.
 XX Modified pro-urokinase has a longer half-life in blood, and dissolves
 CC thrombin without causing the spontaneous bleeding associated with
 CC urokinase. The modification puts an epidermal growth factor domain
 CC into the protein, containing the sequence Asn-X-Ser or Asn-X-Thr
 CC where X is any residue.
 CC Plasmid pHR27 is disclosed as containing the modified sequence.
 XX Sequence 411 AA;
 Query Match 100.0%; Score 1508; DB 11; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVYCGSLIS 60
 DB 136 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAAIYRRHRGGSVTVYCGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCETIGFGKENSVDLYPEQLKMT 180
 DB 256 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFTSCETIGFGKENSVDLYPEQLKMT 315
 QY 181 VVKLISHRECOQPHYYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 316 VVKLSHRECOQPHYGVSEVTTKMLCAADPOKWTDSQCQSGPLVCSLQGRMTLTGIVS 375
Qy 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
Db 376 WGRGKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 38
AAR34584
ID AAR34584 standard; Protein; 411 AA.
AC AAR34584;
XX
DT 25-MAR-2003 (updated)
DT 14-SEP-1993 (first entry)
XX
DE Mutant human prourokinase.
XX
XX pUK; increased half life; improved fibrin affinity.
XX
OS Homo sapiens.
XX
PN EP541952-A1.
XX
PD 19-MAY-1993.
XX
PP 06-OCT-1992; 92EP-0117000.
XX
PR 07-OCT-1991; 91JP-0289257.
XX
XX (GREC) GREEN CROSS CORP.
XX
XX Amatsuji Y, Hirose M, Morita M, Tanabe T;
XX
XX WPI; 1993-160551/20.
XX
XX N-PSDB; AA041450.
XX
XX New human pro-urokinase mutants with thrombolytic activity - have
XX a neutral aminoacid in the epidermal growth factor region replaced
XX with a basic aminoacid, or an acid residue replaced with a non-
XX acidic residue
XX
XX Claim 1; Page 17-20; 38pp; English.
XX
XX The sequence is that of a mutant human prourokinase, in which a
XX neutral amino acid in the epidermal growth region has been replaced
XX by a basic amino acid, or an acidic amino acid has been replaced
XX by a non-acidic amino acid. Preferred replacements are 16Gly->Lys,
XX 38Gly->Lys and 45Asp->Asn. The mutant has an increased half-life in
XX blood as compared to the prior art mutant with a deleted EGF region.
XX It has improved affinity for fibrin, and has other features the same
XX as human prourokinase.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 14; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.2e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPEELKFCQGGKTLRPREFKIGGFTTLENQWFAAIYRRHRGGSVTVCGSLIS 60
Db 136 KPSSPEELKFCQGGKTLRPREFKIGGFTTLENQWFAAIYRRHRGGSVTVCGSLIS 195
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSQTQEMKFEVENLILHKDYSADTLAHND 120
Db 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSQTQEMKFEVENLILHKDYSADTLAHND 255
Qy 121 IALLKIRSKGRCAPSRITQICLPMSVNDPQGTSCITGFGKENSVDLYLPEQLKMT 180
Db 256 IALLKIRSKGRCAPSRITQICLPMSVNDPQGTSCITGFGKENSVDLYLPEQLKMT 315

Qy 181 VVKLSHRECOQPHYGVSEVTTKMLCAADPOKWTDSQCQSGPLVCSLQGRMTLTGIVS 240
Db 316 VVKLSHRECOQPHYGVSEVTTKMLCAADPOKWTDSQCQSGPLVCSLQGRMTLTGIVS 375
Qy 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
Db 376 WGRGKALDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 39
AAR62991
ID AAR62991 standard; protein; 411 AA.
AC AAR62991;
XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis.
XX
OS Homo sapiens.
XX
FH Key
FT Domain
FT -Location/Qualifiers
FT 297..313
FT /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
XX W09501427-A1.
XX
XX 12-JAN-1995.
XX
XX 28-JUN-1994; 94WO-US07278.
XX
XX 02-JUL-1993; 93US-0087163.
XX
XX (NEWEM-) NEW ENGLAND DEACONESS HOSPITAL.
XX
XX Gurewicz V, Liu J;
XX
XX WPI; 1995-060991/08.
XX
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX
XX Disclosure; Fig 1; 46pp; English.
XX
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.2e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY 1 KPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVYVCGGSLIS 60
Db 136 KPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVYVCGGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
Db 196 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 255
QY 121 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 180
Db 256 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 315
QY 181 VVKLIISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 240
Db 316 VVKLIISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGCALKDKPGVYTVRSHFLPWIRSHTKENGAL 276
Db 376 WGRGCALKDKPGVYTVRSHFLPWIRSHTKENGAL 411

RESULT 40
AAV39343
ID AAY39343 standard; protein; 411 AA.
AC AAY39343;
XX
XX 01-DEC-1999 (first entry)
XX Human pro-urokinase.
XX Serine protease; plasminogen; plasmin; activation; matrix; cancer;
XX tumour; metastasis; X-ray crystallography; inhibitor.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..158
FT Disulfide-bond /label= Mature_urokinase_A_chain
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Cleavage-site 135..136
FT /note= "Additional cleavage at this site generates low
FT molecular weight (LMW) urokinase"
FT Protein 1..135
FT Disulfide-bond /label= Low_molecular_weight_urokinase
FT Disulfide-bond 148..279
FT Cleavage-site /note= "Links mature urokinase A- and B-chains"
FT 158..159
FT /note= "Cleavage at this site generates mature urokinase
FT A- and B-chains"
FT Protein 159..411
FT Disulfide-bond /label= Mature_urokinase_B_chain
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
FT Modified-site 302
FT /note= "N-glycosylated"
FT Cleavage-site 405..406
XX WO9945379-A2.
XX
XX 10-SEP-1999.
XX
XX 05-MAR-1999; 99WO-US04967.
XX
```

```
PR 06-MAR-1998; 98US-0036184.
XX (ABBO ) ABBOTT LAB.
XX
XX Nienaber VL, Greer J, Abad-zapatero C, Norbeck DW;
XX WPI; 1999-571607/48.
XX Identifying ligands for target biomolecules using X-ray crystallography
XX
XX Example 1; Fig 5; 57pp; English.
XX
XX This sequence represents human pro-urokinase. The mature urokinase
XX consists of an A- and B-chain, linked by a single disulphide bond, and
XX is generated by proteolytic cleavage of the peptide bond between Lys 158
XX and Ile 159. Additional cleavage of the peptide bond between Lys 135
XX and Lys 136 generates a low molecular weight urokinase. The urokinase
XX A-chain contains an EGF-like domain and a kringle domain, while the
XX B-chain contains the catalytic domain. Urokinase is a serine protease
XX and is strongly associated with tumour cells. Urokinase activates
XX plasminogen which, in turn, activates the matrix metalloproteinases.
XX Plasmin and the metalloproteinases degrade the extracellular matrix and
XX promote tumour growth and metastasis. Inhibitors that specifically target
XX urokinase may serve as effective anticancer agents. A novel method for
XX identifying such ligands used X-ray crystallography to determine if a
XX complex is formed between a ligand and a target biomolecule. However,
XX crystals of a native urokinase/inhibitor complex had poor diffraction
XX quality. Human urokinase was therefore engineered so that it would
XX produce crystals with the desired qualities. This engineered urokinase
XX was designated mu-UK (AAY39344).
XX
XX Sequence 411 AA;
XX
XX Query Match 100.0%; Score 1508; DB 20; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVYVCGGSLIS 60
Db 136 KPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVYVCGGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
Db 196 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHND 255
QY 121 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 180
Db 256 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 315
QY 181 VVKLIISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 240
Db 316 VVKLIISHRECOOPHYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGCALKDKPGVYTVRSHFLPWIRSHTKENGAL 276
Db 376 WGRGCALKDKPGVYTVRSHFLPWIRSHTKENGAL 411

RESULT 41
AAV42284
ID AAY42284 standard; protein; 411 AA.
XX
XX AAY42284;
XX
XX 01-DEC-1999 (first entry)
XX Human pro-urokinase.
XX Serine protease; plasminogen; plasmin; activation; matrix; cancer;
XX tumour; metastasis; X-ray crystallography; inhibitor.
XX Homo sapiens.
XX
```

XX PH Key Location/Qualifiers
 FT Protein 1..158
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Cleavage-site 135..136
 FT /note= "Additional cleavage at this site generates low molecular weight (LMW) urokinase"
 FT Protein 1..135
 FT /label= Low_molecular_weight_urokinase
 FT Disulfide-bond 148..279
 FT /note= "Links mature urokinase A- and B-chains"
 FT Cleavage-site 158..159
 FT /note= "Cleavage at this site generates mature urokinase A- and B-chains"
 FT Protein 159..411
 FT /label= Mature_urokinase_B_chain
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 FT Modified-site 302
 FT /note= "N-glycosylated"
 FT Cleavage-site 405..406
 XX WO9945389-A2.
 XX 10-SEP-1999.
 XX 01-MAR-1999; 99WO-US04518.
 XX 06-MAR-1998; 98US-0036184.
 XX (ABBO) ABBOTT LAB.
 XX Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;
 WPI; 1999-551079/46.
 XX Identifying ligands for target biomolecules using X-ray
 FT crystallography, used for designing ligands with improved biological
 FT activity for target receptor -
 XX Example 1; Fig 5; 57pp; English.
 XX This sequence represents human pro-urokinase. The mature urokinase
 CC consists of an A- and B-chain, linked by a single disulphide bond, and
 CC is generated by proteolytic cleavage of the peptide bond between Lys 158
 CC and Ile 159. Additional cleavage of the peptide bond between Lys 135
 CC and Lys 136 generates a low molecular weight urokinase. The urokinase
 CC A-chain contains an EGF-like domain and a kringle domain, while the
 CC B-chain contains the catalytic domain. Urokinase is a serine protease
 CC and is strongly associated with tumour cells. Urokinase activates
 CC plasminogen which, in turn, activates the matrix metalloproteinases.
 CC Plasmin and the metalloproteinases degrade the extracellular matrix and
 CC promote tumour growth and metastasis. Inhibitors that specifically target
 CC urokinase may serve as effective anticancer agents. A novel method for
 CC identifying such ligands used X-ray crystallography to determine if a
 CC complex is formed between a ligand and a target biomolecule. However,
 CC crystals of a native urokinase/inhibitor complex had poor diffraction
 CC quality. Human urokinase was therefore engineered so that it would
 CC produce crystals with the desired qualities. This engineered urokinase
 CC was designated mu-UK (AA42285).
 XX Sequence 411 AA;
 Query Match 100.0%; Score 1508; DB 20; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPBELKFOCGQKTLRPRFKIIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60
 DB 136 KPSSPPBELKFOCGQKTLRPRFKIIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQCEMKFEVENLILHKDYSADTTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQCEMKFEVENLILHKDYSADTTLAHND 255
 QY 121 IALLKIRSEKGRCAQPSRTIQICLPSPMYNDPQFTSCETGFGKENSTDYLYPQLKMT 180
 DB 256 IALLKIRSEKGRCAQPSRTIQICLPSPMYNDPQFTSCETGFGKENSTDYLYPQLKMT 315
 QY 181 VVKLISHRECQOPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 316 VVKLISHRECQOPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 376 WGRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
 RESULT 42
 AAY92836
 ID AAY92836 standard; Protein; 411 AA.
 XX
 AC AAY92836;
 XX
 DT 29-AUG-2000 (first entry)
 XX
 DE Urokinase plasminogen activator (uPA).
 XX N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;
 KW anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;
 KW anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;
 KW anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;
 KW thrombolytic.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX WO200026353-A1.
 PD 11-MAY-2000.
 XX 28-OCT-1999; 99WO-US25210.
 XX 29-OCT-1998; 98US-0181816.
 XX (ANGS-) ANGSTROM PHARM INC.
 XX Mazar AP, Jones TR;
 XX WPI; 2000-365605/31.
 XX New cyclic peptide, useful for treatment or diagnosis of e.g. tumors
 PT and other diseases involving cell proliferation or migration, targets
 PT the urokinase plasminogen activator receptor

XX Disclosure; Fig 1; 93pp; English.

PS The present sequence shows the wild-type urokinase plasminogen activator

XX (uPA). Cyclic peptides based on the amino acids residues 20-30 (the

CC receptor-binding region) of uPA are claimed. These cyclic peptides target

CC the uPA receptor (UPAR), allowing therapeutic or diagnostic agents to be

CC delivered to UPAR-expressing cells. The cyclic peptides are used,

CC optionally when linked to a therapeutic agent, to inhibit migration,

CC invasion and proliferation of cells, or angiogenesis, or to induce

CC apoptosis. Particularly they are used, in human or veterinary medicine,

CC to treat diseases characterized by these processes, e.g. solid tumors,

CC leukaemia or lymphoma (or their metastases); benign hyperplasia;

CC atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular

CC glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc.,

CC most particularly growth, invasion and metastasis of tumors. When

CC labeled, the cyclic peptides can be used for diagnostic detection of UPAR

CC (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and

CC when immobilized they are used to isolate uPA or cells that express

CC them. The cyclic peptides are stable, soluble in water, bind strongly to

CC uPA, are relatively inexpensive to produce and may be derivatized by

CC attachment of therapeutic or diagnostic agents without significantly

CC affecting their binding. Since they target uPA, they should have

CC relatively low systemic toxicity and only low doses are required.

XX

SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 21; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.2e-125;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPELKKQCGQKTLRPFKLIIGGFTTIENQPFPAALYRHRGGSVTVYCGGLIS 60

Db 136 KPSSPPELKKQCGQKTLRPFKLIIGGFTTIENQPFPAALYRHRGGSVTVYCGGLIS 195

QY 61 PCWVISAHCFIDYKPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120

Db 196 PCWVISAHCFIDYKPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 255

QY 121 IALLKIRSEKRCQAPSTIOTICLPMSYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 180

Db 256 IALLKIRSEKRCQAPSTIOTICLPMSYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 315

QY 181 VVKLIHRECCQPHYGVSEVTTKMLCAADPWKTDSCQDSGGPLVCSLQGRMTLTGIVS 240

Db 316 VVKLIHRECCQPHYGVSEVTTKMLCAADPWKTDSCQDSGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGKALXDPGVYTRVSHFLPWIRSHTKKEENGLAL 276

Db 376 WGRGKALXDPGVYTRVSHFLPWIRSHTKKEENGLAL 411

RESULT 43

AA20489

ID AAB20489 standard; Protein; 411 AA.

XX

XX AAB20489;

XX

XX 21-JUN-2001 (first entry)

XX

XX Human pro-urokinase plasminogen activator.

DE

XX

KW Urokinase plasminogen activator; uPA; human; tumour;

KW cell migration; cell invasion; cell proliferation; angiogenesis;

KW apoptosis; antitumour; diagnosis; therapy.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FH Disulfide-bond 11..19

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Cleavage-site 158..159

FT /note= "Cleavage at this site results in the

FT formation of the two-chain active uPA

FT (tcuPA)"

FT

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

FT Domain 1..135

FT /label= ATF

FT /note= "the ATF domain alternatively comprises

FT residues 1-43"

FT

FT Domain 4..43

FT /label= Growth_factor_domain

XX

XX WO200125410-A2.

XX

XX 12-APR-2001.

XX

XX 27-SEP-2000; 2000WO-US26502.

XX

XX 01-OCT-1999; 99US-0157012.

XX

XX (ANGS-) ANGSTROM PHARM INC.

XX

XX Mazar AP, Jones TR;

XX

XX WPI; 2001-290611/30.

DR

XX

XX Novel urokinase plasminogen activator cell surface receptor-targeting

PT protein or peptide, useful for inhibiting angiogenesis or cell

PT migration, invasion or proliferation, is diagnostically or

PT therapeutically labeled -

XX

PS Disclosure; Fig 1; 35pp; English.

XX

CC The present sequence is that of human pro-urokinase plasminogen

CC activator (pro-uPA). The invention provides a uPA receptor (UPAR)

CC targeting protein or peptide that is labelled and used in methods of

CC diagnosis and therapy. The labelled protein or peptide preferably

CC has the following properties: it comprises at least 38 amino acid

CC residues, including residues 13-30 of the UPAR binding site of uPA;

CC competes with labelled DFP-uPA for binding to a cell or molecule

CC that has a binding site for uPA; has an IC50 value of about 10 nM

CC or less; and is not a fusion protein. Preferred molecules are uPA,

CC (residues 1-411), single chain uPA, tcuPA (inactivated with the

CC suicide inhibitor diisopropyl fluorophosphate), the N-terminal ATF

CC fragment (amino acids 1-135 or 1-143) of uPA, or the growth factor

CC domain (residues 4-43). Suitable labels include a radionuclide,

CC a PET-imageable agent, an MRI-imageable agent, a fluororescer, a

CC fluorogen, a chromophore, a chromogen, a phosphorescer, a

CC chemiluminescer or a bioluminescer. The methods are used to

CC inhibit cell migration, cell invasion (preferably invasiveness of

CC tumour cells), cell proliferation or angiogenesis, or to induce

CC apoptosis, preferably in the treatment of a subject having a disease

CC or condition associated with undesired cell migration, invasion,

CC proliferation or angiogenesis (claimed). The protein or peptide is

CC also useful for treating diseases or conditions including primary

CC growth of a solid tumour, leukaemia or lymphoma, tumour invasion,

CC metastasis, atherosclerosis, myocardial angiogenesis, telangiectasia,

CC corneal disease, rubecosis, neovascular glaucoma, diabetic and other

CC retinopathy, macular degeneration, arthritis, fibrosis, wound

CC healing with scarring and fibrosis, peptic ulcers, bone fracture,

CC keloids, or a disorder of vasculogenesis, haematopoiesis, ovulation,

CC menstruation, pregnancy or placental associated with pathogenic

CC cell invasion or with angiogenesis. The protein or peptide probe is

CC internalised by the cells to which it binds, e.g. tumour cells, and

CC is useful for imaging techniques in which it reduces the background

CC signal relative to specifically bound probes. This uptake permits

CC clearance of circulating probe so that the ratio of labelled probe
 CC inside tumour cells to the probe elsewhere in the body increases.

XX
 SQ Sequence 411 AA;
 Query Match 100.0%; Score 1508; DB 22; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVYCGGSLIS 60
 DB 136 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVYCGGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180
 DB 256 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 315
 QY 181 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTGIVS 240
 DB 316 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 376 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 44
 AAB74797
 ID AAB74797 standard; Protein; 411 AA.
 AC AAB74797;
 XX 12-JUN-2001 (first entry)
 XX Prourokinase protein sequence.
 KW Prourokinase; Pro-309; mutagenic; urokinase; zymogen; mutant;
 KW lowered fibrinogen dissolving activity; fibrin; E segment; D segment;
 KW lowered non-specific fibrin dissolving zymogen activation.
 XX Unidentified.

XX Key Location/Qualifiers
 FH Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380
 XX CN1277262-A.
 PN 20-DEC-2000.
 PD 10-JUL-2000; 2000CN-0109829.
 PF 10-JUL-2000; 2000CN-0109829.
 PR (LIUJ/) LIU J.
 PA Sun Z, Liu J;
 XX WPI; 2001-266614/28.

XX Urokinase zymogen mutant -
 XX Example; Fig 1; lipp; Chinese.

XX The present invention describes a prourokinase mutant comprising the
 CC amino acid sequence point mutation at proline 309. The mutation makes
 CC the mutant have an intrinsic activity 2.5-20 times lower than that of
 CC natural prourokinase, including lowered fibrinogen dissolving activity
 CC and lowered non-specific fibrin dissolving zymogen activation. Compared
 CC with the natural prourokinase, the fibrin dissolving zymogen activation
 CC of the mutant may be promoted by not only the E segment of degraded
 CC fibrin but also the D segment. The present sequence represents a
 CC wild type prourokinase protein sequence which is used in an example
 CC from the present invention.
 CC N.B. The sequence in the specification is of poor quality so the
 CC sequence given here is of the indexers best interpretation.

XX SQ Sequence 411 AA;
 Query Match 100.0%; Score 1508; DB 22; Length 411;
 Best Local Similarity 100.0%; Pred. No. 2.2e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVYCGGSLIS 60
 DB 136 KPSSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVYCGGSLIS 195
 QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 196 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
 QY 121 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180
 DB 256 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 315
 QY 181 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTGIVS 240
 DB 316 VVKLISHRECQPHYGGSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTGIVS 375
 QY 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 376 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 45
 AAE16544
 ID AAE16544 standard; Protein; 411 AA.
 AC AAE16544;
 XX 09-APR-2002 (first entry)
 XX Human urokinase-type plasminogen activator tcuPA and scuPA protein.
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW male impotence; adult respiratory distress syndrome; tcuPA; scuPA;
 KW two chain urokinase; single chain urokinase.
 OS Homo sapiens.
 XX WO200197752-A2.
 PN 27-DEC-2001.
 PD 13-JUN-2001; 2001WO-US18976.
 PF 20-JUN-2000; 2000US-212874P.
 PR

(UYPE-) UNIV PENNSYLVANIA.

PA

Cines DB, Higazi AA;

XX

WPI; 2002-122240/16.

XX

N-PSDB; AAD27077.

XX

Composition for modulating muscle cell and tissue contractility for treating atherosclerosis, asthma, hypertension, glaucoma, impotence, comprising domains from urokinase-type plasminogen activator -

XX

Claim 9; Fig 1C; 117pp; English.

XX

The invention relates to a composition comprising one or more domains of urokinase-type plasminogen activator (uPA). The composition is used to modulate the contractility and angiogenic activity of a mammalian muscle, endothelial cell or tissue. The composition is used for treating stroke, hypotension, hypertension, atherosclerosis, heart attack, microvascular occlusions, thrombotic microangiopathies, surgically induced thrombotic disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma, diabetic retinopathy, wound healing, clotting disorder, uterine contraction disorder, male impotence, respiratory disease or condition such as asthma, adult respiratory distress syndrome, primary pulmonary hypertension, microvascular thrombotic occlusion, and a disorder associated with chronic intrapulmonary fibrin formation. The present sequence is human urokinase-type plasminogen activator (uPA) two chain urokinase (tcuPA) and single chain urokinase (scuPA) protein.

Sequence 411 AA;

XX

Query Match 100.0%; Score 1508; DB 23; Length 411;

XX

Best Local Similarity 100.0%; Pred. No. 2.2e-125;

XX

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

1 KPSSPPEELKFCQGGKTLRPRFKIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60

QY

136 KPSSPPEELKFCQGGKTLRPRFKIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 195

DB

61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120

QY

196 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 255

DB

121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPFGTSCEITGFGKENSTDYLYPEQLKWT 180

QY

256 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPFGTSCEITGFGKENSTDYLYPEQLKWT 315

DB

181 VKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

QY

316 VKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375

DB

241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

QY

376 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

DB

RESULT 46

XX

AAR10172

XX

AAR10172 standard; Protein; 421 AA.

XX

AAR10172;

XX

25-MAR-2003 (updated)

XX

21-MAR-1991 (first entry)

XX

Human prourokinase derivative AFUK (k.k).

XX

human prourokinase; blood coagulation Factor XIII; thrombolysis;

XX

Homo sapiens.

XX

WO9015867-A.

XX

Key

XX

Location/Qualifiers

XX

Key

XX

Location/Qualifiers

XX

Key

XX

Location/Qualifiers

XX

Key

XX

Location/Qualifiers

XX

Key

XX

27-DEC-1990.

PD

07-JUN-1990; 90WO-JP00742.

XX

15-MAY-1990; 90JP-0123163.

XX

13-JUN-1989; 89JP-0150161.

XX

(NIPS) NIPPON SODA CO.

XX

Kobayashi Y, Watabe K, Mukohara Y, Satoh M, Nakamura H;

XX

WPI; 1991-022232/03.

XX

N-PSDB; AAQ10224.

XX

Human prourokinase-like polypeptide(s) with thrombolytic action

XX

- have oligopeptide(s) which bind covalently to blood

XX

clot(thrombus) by action of factor XIII

XX

Disclosure; Fig 7; 44pp; English.

XX

APUK(k.k) is an example of a prourokinase derivative which has increased affinity for blood clots and attacks the components of a thrombus whilst sparing the circulating clotting proteins and platelets. The first 13 N-terminal amino acids replace the first 3 amino acids of natural urokinase. This derivative has a fibrin

XX

lysis activity of 9.2 x 10 power4 international units/O.D.280.

XX

See also AAQ10222-3 and AAQ10225.

XX

(Updated on 25-MAR-2003 to correct PA field.)

XX

Sequence 421 AA;

XX

Query Match 100.0%; Score 1508; DB 12; Length 421;

XX

Best Local Similarity 100.0%; Pred. No. 2.3e-125;

XX

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

1 KPSSPPEELKFCQGGKTLRPRFKIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60

QY

146 KPSSPPEELKFCQGGKTLRPRFKIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 205

DB

61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120

QY

206 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 265

DB

121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPFGTSCEITGFGKENSTDYLYPEQLKWT 180

QY

266 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPFGTSCEITGFGKENSTDYLYPEQLKWT 325

DB

181 VKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

QY

326 VKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 385

DB

241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

QY

386 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 421

DB

RESULT 47

XX

AAR50114

XX

AAR50114 standard; Protein; 431 AA.

XX

AAR50114;

XX

27-SEP-1991 (first entry)

XX

Sequence encoded by the signal sequence and noncoding region of the

XX

pro-UK structural gene (Sequence II).

XX

Enzyme; thrombolysis therapy; embolic disease;

XX

single-chain pro-urokinase.

XX

Homo sapiens.

XX

Key

XX

Location/Qualifiers

XX

Key

XX

Location/Qualifiers

XX

Key

XX

Location/Qualifiers

XX

Key

XX

Location/Qualifiers

XX

Key

XX

FT Domain 21..177
 FT /label= A chain
 FT 179..431
 FT /label= B chain
 FT Peptide 1..20
 FT /label= signal peptide

XX EPI54272-A.
 XX 11-SEP-1985.
 XX 23-FEB-1985; 85EP-0102031.
 XX 31-JAN-1985; 85JP-0017969.
 XX 27-FEB-1984; 84JP-0037119.
 XX (GREG) GREEN CROSS CORP.
 XX Hiramatsu R, Kaneda T, Nagai M, Arimura H, Nishida M;
 XX Suyama T;
 XX WPI; 1985-224693/37.
 XX N-PSDB; AAN50138.

XX Glycosylated single-chain pro-urokinase - prep. by cultivating
 XX animal cells transformed by DNA prep. from m RNA
 XX Disclosure; Page 8-10; 64pp; English.

XX The inventors claim a method of producing single-chain
 XX pro-urokinase by using as template, mRNA obtd. from cells of an
 XX established human kidney-derived cell line. The urokinase is used to
 XX treat thrombosis and embolic diseases as well as in the treatment of
 XX diseases in combination with anticancer agents.

XX Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 6; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVYCGSGLIS 60
 DB 156 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVYCGSGLIS 215
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
 QY 121 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQFTSCETITGFKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQFTSCETITGFKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYYSSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYYSSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 48
 AAP60674
 ID AAP60674 standard; Protein; 431 AA.
 XX AAP60674;
 XX 25-MAR-2003 (updated)
 DT 16-OCT-1991 (first entry)
 XX Modified human pro-urokinase clone.
 XX

KW Fibrinolytic; protease; trypsin; thrombin; plasmin.

XX Homo sapiens.

XX MO8604351-A.

XX 31-JUL-1986.

XX 24-JAN-1986; 86WO-JP00031.

XX 20-JUL-1985; 85JP-0159294.

XX 25-JAN-1985; 85JP-0011033.

XX 20-APR-1985; 85JP-0083611.

XX (SAGA) SAGAMI CHEM RES CENTRE.

XX (MIYA/) MIYAKE T.

XX (HODO) HODOGAYA-CHEM CO LTD.

XX (NIPS) NIPPON SODA CO.

XX (TOYJ) TOYO SODA MFG CO LTD.

XX (NISC) NISSAN CHEM IND LTD.

XX (NIPC) NIPPON CHEM IND CO LTD.

XX MIYAKE T, HIBINO Y, KOBAYASHI Y, WATABE K, OMORI M, MIKI T;

PI YOKOYAMA M, MATSUMOTO R;

XX WPI; 1986-212124/32.

XX N-PSDB; AAN60591.

XX Human pro-urokinase stable to protease and 135,157-modified - prep.
 XX using transformant e.coli and having long-acting fibrinolytic activity
 XX Disclosure; Fig 4; 98pp; Japanese.

XX Sequence is modified from the human pro-urokinase with a non-basic
 XX residue replacing Lys135. The product may be expressed from a
 XX transformed host such as E.coli, it has long-acting fibrinolytic
 XX activity and is stable to proteases such as trypsin, thrombin and
 XX plasmin.
 XX (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 7; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVYCGSGLIS 60
 DB 156 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVYCGSGLIS 215
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
 QY 121 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQFTSCETITGFKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQFTSCETITGFKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYYSSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYYSSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 49
 AAP60783
 ID AAP60783 standard; Protein; 431 AA.
 XX AAP60783;
 XX

DT	25-MAR-2003 (updated)	XX	25-MAR-2003 (updated)	XX
DT	23-OCT-1991 (first entry)	DT	19-MAY-1991 (first entry)	DT
DE	Human urokinase.	XX	Sequence of human prourokinase and leader.	XX
XX	E.coli; high molecular urokinase.	XX	Cardiovascular disease treatment; fibrin affinity; thrombolytic; enzyme; protease.	XX
XX	Homo sapiens.	OS	Homo sapiens.	OS
PH	Key	XX	Location/Qualifiers	XX
FT	Protein	FT	Peptide	FT
XX	JP61181377-A.	XX	1..20	XX
XX	14-AUG-1986.	XX	/label= leader	XX
XX	25-JAN-1985; 85JP-0011032.	XX	21..431	XX
XX	25-JAN-1985; 85JP-0011032.	XX	/label= prourokinase	XX
XX	(NISC) NISSAN CHEM IND LTD.	PN	EP231883-A.	PN
XX	(HODO) HODOCAYA CHEM IND CO LTD.	XX	12-AUG-1987.	XX
XX	(SAGA) SAGAMI CHEM RES CENTRE.	XX	29-JAN-1987; 87EP-0101209.	XX
XX	(CENG) CENTRAL GLASS CO LTD.	XX	31-JAN-1986; 86JP-0017734.	XX
XX	(NIPS) NIPPON SODA CO.	PR	30-JAN-1987; 87JP-0018626.	PR
XX	(TOYJ) TOYO SODA MFG CO LTD.	XX	(SAGA) SAGAMI CHEM RES CENTRE.	XX
XX	(NISC) NISSAN CHEM IND LTD.	PA	(NIPS) NIPPON SODA CO.	PA
XX	(TOYJ) TOYO SODA MFG CO LTD.	PA	(CENG) CENTRAL GLASS CO LTD.	PA
XX	WPI; 1986-254744/39.	PA	(TOYJ) TOYO SODA MFG CO LTD.	PA
XX	N-PSDB; AAN60703.	PA	(NISC) NISSAN CHEM IND LTD.	PA
XX	Human urokinase gene - has N-end of aminoacid sequence coded by codon used in Escherichia coli.	PA	(NISC) NISSAN CHEMICAL INDS KK.	PA
XX	Disclosure; Fig 2; 19pp; Japanese.	XX	Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;	XX
XX	The claimed gene product may be expressed in a transformed E.coli host, for the efficient production of high molecular human urokinase.	PI	WPI; 1987-222882/32.	PI
XX	The N-terminal of the protein expressed by the transforming plasmid is replaced with a codon frequently used in E.coli.	DR	N-PSDB; AAN70390.	DR
XX	(Updated on 25-MAR-2003 to correct PA field.)	XX	Hybrid plasminogen activator-like polypeptide - having a region for affinity to fibrin from tissue plasminogen activator and a region from prourokinase	XX
XX	Sequence 431 AA;	PS	Disclosure; Fig 2(1-5); 64pp; English.	PS
XX	Query Match	XX	The TPA portion of the claimed hybrid polypeptide (see FT) may consist of 2 kringles from N-terminal first serine to 219th glycine of human TPA, 1 kringles from 128th serine to 219th glycine of human TPA or half a kringles from 161st methionine to 219th glycine (see AAP70257). The C-terminal half of the hybrid polypeptide may contain an AA sequence from 150th glutamine to C-terminal 41th leucine of prourokinase (see AAP70258).	XX
XX	Best Local Similarity 100.0%; Score 1508; DB 7; Length 431;	CC	(Updated on 25-MAR-2003 to correct PA field.)	CC
XX	Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	Sequence 431 AA;	XX
QY	1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60	QY	1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60	QY
DB	156 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 215	DB	156 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 215	DB
QY	61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120	QY	61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120	QY
DB	216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275	DB	216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275	DB
QY	121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 180	QY	121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 180	QY
DB	276 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 335	DB	276 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 335	DB
QY	181 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240	QY	181 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240	QY
DB	336 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395	DB	336 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395	DB
QY	241 WGRGCALKOKPGVYTRVSHFLPWIRSHTRKEENGLAL 276	QY	241 WGRGCALKOKPGVYTRVSHFLPWIRSHTRKEENGLAL 276	QY
DB	396 WGRGCALKOKPGVYTRVSHFLPWIRSHTRKEENGLAL 431	DB	396 WGRGCALKOKPGVYTRVSHFLPWIRSHTRKEENGLAL 431	DB
RESULT 50				
AAP70258				
ID	AAP70258 standard; Protein; 431 AA.			
XX				
AC	AAP70258;			

Db 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWTDSCQDGGPLVCSLQGRMTLTGIVS 395

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 51

ID AAP70250 standard; protein; 431 AA.

AC AAP70250;

XX 21-FEB-1991 (first entry)

DT Sequence of A- and B-chains of high molecular weight-type urokinase

DE obtained from human urine.

DE Enzyme; protease; thrombolytic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note="Signal"

XX 21..431

PN EP232544-A.

XX 19-AUG-1987.

XX 24-DEC-1986; 86EP-0118034.

XX 25-DEC-1985; 85JP-0290325.

XX (GREC) GREEN CROSS CORP.

XX Kaneda T, Okabayashi K, Hayasuke N, Hiramatsu R, Nagai M;

PI Arimura H;

XX WPI; 1987-229483/33.

DR N-PSDB; AAN70356.

XX Prodn., of pure natural poly:peptide(s) and protein(s) -

PT comprises culturing transformed human kidney cells in medium

PT free from bovine serum

XX Example; Pages 12-15; 30pp; English.

XX Urokinase cDNA (see AAN70356) was inserted in a vector, and a dominant

CC selection marker contg. plasmid DNA was prepd. The vector is

CC pSV-G(l)-Neo(r), which is pSV-G(l) contg. the tns-derived neomycin-

CC resistant gene. The vector was used for DNA transfection of the

CC recipient human foetal cells. Selected transformed cells were

CC cultured etc. The human prourokinase cDNA was normally transcribed

CC to give the glycoprotein.

XX Sequence 431 AA;

SQ

Query Match 100.0%; Score 1508; DB 8; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.4e-125;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPEELKFCGGKTLPRFKIIGGETTIENOPFAAIVRRHGGSVTVVCGSLIS 60

Db 156 KPSSPEELKFCGGKTLPRFKIIGGETTIENOPFAAIVRRHGGSVTVVCGSLIS 215

Qy 61 PCWVISATCFIDYPKKEDIIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120

Db 216 PCWVISATCFIDYPKKEDIIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

Qy 121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180

Db 276 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 335

Qy 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWTDSCQDGGPLVCSLQGRMTLTGIVS 395

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 52

ID AAP80430 standard; protein; 431 AA.

AC AAP80430;

XX 25-MAR-2003 (updated)

DT 14-SEP-1990 (first entry)

XX Deduced AA sequence of the single chain urokinase plasminogen activator

DE (SCU-PA) cDNA insert prepared from human Hep3 cells.

XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;

KW glyceraldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;

KW thrombosis treatment.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein 21..411

XX EP288435-A.

XX 26-OCT-1988.

XX 11-APR-1988; 88EP-0810234.

XX 15-APR-1987; 87GB-0009081.

PR 16-JUN-1987; 87GB-0014059.

PR 04-DEC-1987; 87IE-0003299.

XX (CIBA) CIBA GEIGY AG.

PA Meyhack B, Heim J, Burgi R;

XX WPI; 1988-301440/43.

DR N-PSDB; AAN80981.

XX Prodn. of human single chain urokinase type plasminogen activator -

PT by culturing yeast strain transformed with hybrid vector contg.

PT yeast expression control sequences

XX Example 1; Fig 2; 48pp; English.

XX The patent is for the prodn. of human single chain urokinase-type

CC plasminogen activator (UTPA). Mutants of scu-PA are especially those

CC which render the protein protease resistant. Such scu-PA mutants are

CC covalently modified at sites of proteolysis by proteases occurring in

CC blood such as thrombin or plasmin, so that they are no longer

CC susceptible to protease hydrolysis at these locations. The target sites

CC include Lys135 to Lys136 (cleavage at this site generates the so-called

CC low molecular weight form of scu-PA or LUK), Arg156 to Phe157

CC (susceptible to thrombin attack) and Lys158 to Ile159 (cleavage at this

CC site by plasmin generates tuc-PA). Suitable scu-PA mutants have site

CC specific substitutions, insertions or deletions of residues at

CC one or more of these target sites. Especially preferred are those

CC mutants in which one amino acid residue or both amino acid residues

CC forming the target sites are deleted or in which at least one of these

CC amino acid residues is replaced by another amino acid residue so that

CC the resulting mutants are resistant to proteolytic attack. The UTPA

CC proteins exhibit the biological activity of natural human UTPA without

any refolding procedure being necessary. They can be used as for known
 PAs in humans for the prevention or treatment of thrombosis or other
 conditions where it is desired to produce local fibrinolytic or
 proteolytic activity.
 (Updated on 25-MAR-2003 to correct PR field.)

Sequence 431 AA;
 Query Match 100.0%; Score 1508; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAALYRHRGGSVTVYVCGSGLIS 60
 DB 156 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAALYRHRGGSVTVYVCGSGLIS 215
 QY 61 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHHND 120
 DB 216 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHHND 275
 QY 121 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYVSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYVSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGAL 431

RESULT 53
 AAP81204
 ID AAP81204 standard; protein; 431 AA.
 XX
 AC AAP81204;
 DT 25-MAR-2003 (updated)
 DT 03-DEC-1990 (first entry)
 XX
 DE Pro-urokinase with signal sequence.
 XX
 KW pro-urokinase (pro-UK); plasminogen activator; pUK33; ss.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..20
 FT /label=signal peptide
 FT Protein 21..431
 FT /label=pro-urokinase
 XX
 PN EP265874-A.
 XX
 XX 04-MAY-1988.
 XX
 PF 23-OCT-1987; 87EP-0115600.
 XX
 PR 23-OCT-1986; 86JP-0253078.
 XX
 PA (GREC) GREEN CROSS CORP.
 XX
 XX Amatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;
 PI WPI; 1988-121000/18.
 XX
 DR N-PSDB; AAN81558.
 XX
 XX Glycosylated single-chain pro-urokinase prodn -
 PT by cultivating DHFR gene-deficient CHO-K1 cells transformed
 PT with a plasmid contg cDNA, SV40 promoter and DHFR gene.
 XX
 PS Disclosure; Page ?; ?pp; English.
 XX

The Arg at position 2 is encoded by TGA(sic). Possible error in the
 specification. Should read CGA ?
 CC The pro-UK gene was derived from plasmid pUK33. The cDNA was
 CC synthesized using urokinase mRNA isolated from a human kidney cell
 CC line. Pro-UK was cloned into a SV40 promoter-contg. plasmid ,down-
 CC stream of the promoter. This plasmid was then ligated to a DHFR-
 CC gene contg. plasmid so that pro-UK and DHFR are inserted in
 CC opposite directions. The recombinant plasmid was used to transform
 CC CHO-K1 cell derived DHFR gene-deficient host cells to produce
 CC glycosylated single-chain pro-UK.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX

Sequence 431 AA;
 Query Match 100.0%; Score 1508; DB 9; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAALYRHRGGSVTVYVCGSGLIS 60
 DB 156 KPSSPPEELKFCQCGQKTLPRFKIIGGEFTTIENQPFPAALYRHRGGSVTVYVCGSGLIS 215
 QY 61 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHHND 120
 DB 216 PCWISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHHND 275
 QY 121 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRSKRGCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYVSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYVSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGAL 431

RESULT 54
 AAP92119
 ID AAP92119 standard; protein; 431 AA.
 XX
 AC AAP92119;
 DT 25-MAR-2003 (updated)
 DT 29-JUN-1990 (first entry)
 XX
 DE Natural human prourokinase.
 XX
 KW Human prourokinase; antithrombotic; derivative.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note="Optional in new deriv."
 FT Misc-difference 2..155 /note="Incorporated into new deriv."
 FT Misc-difference 135 /note="May be replaced by a non-basic AA in new deriv."
 FT Misc-difference 156 /note="Undefined residue in new deriv."
 FT Misc-difference 157 /note="Pro, Gly, Ala or Val in new deriv."
 FT Misc-difference 158 /note="Lys or Arg in new deriv."
 XX
 XX WO8901513-A.
 PN
 XX 23-FEB-1989.
 PD
 XX 18-AUG-1988; 88WO-JP00815.
 PF

XX 19-AUG-1987; 87JP-0204149.
 XX (SAGA) SAGAMI CHEM RES CENTRE.
 PA (CENG) CENTRAL GLASS CO LTD.
 PA (HODO) HODOGAYA CHEM KK.
 PA (NIPS) NIPPON SODA CO.
 PA (NISC) NISSAN CHEM IND LTD.
 XX Kobayashi Y, Omori M, Yamada C;
 XX WPI; 1989-068869/09.
 DR N-PSDB; AAN91075.
 XX Anthrbotomic fast-acting pro-urokinase deriv. -
 PT Produced by culture of E. coli transformant contg. new plasmid
 PT of PMUT9Q family.
 XX Disclosure; Fig 1; 75pp; Japanese.
 XX A human prourokinase (PU) deriv. is new which is based upon residues
 CC 2-155 of natural human prourokinase. The new deriv. is produced by
 CC E. coli J103/pMUT9Q-RPK in culture. It is a fast-acting drug for
 CC the treatment and prevention of thrombosis.
 CC (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX Sequence 431 AA;
 SQ Query Match 100.0%; Score 1508; DB 10; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPEELKFCQCGQKTLPRFKIIGGFTTIENQPFPAIYRRHRGGSVTVCGGSLIS 60
 DB 156 KPSSPEELKFCQCGQKTLPRFKIIGGFTTIENQPFPAIYRRHRGGSVTVCGGSLIS 215
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 275
 QY 121 IALLKIRKEGRCAPSRITQITCLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRKEGRCAPSRITQITCLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 431
 RESULT 55
 AAR04253
 ID AAR04253 standard; protein; 431 AA.
 XX AAR04253;
 XX 25-MAR-2003 (updated)
 DT 12-SEP-1990 (first entry)
 XX Human pro-urokinase from the cDNA of clone pcUK176.
 XX Non-glycosylated; pro-urokinase; E. coli; P1rp promoter; MS-2 RBS.
 XX Synthetic.
 XX EP365894-A.
 XX 02-MAY-1990.

PF 06-OCT-1989; 89EP-0118586.
 XX 11-OCT-1988; 88GB-0023833.
 XX (FARM) FARMITALIA ERBA SPA CARLO.
 XX Brandazza A, Sarmientos P, Orsini G;
 XX WPI; 1990-133447/18.
 DR N-PSDB; AAQ04107.
 XX Non-glycosylated pro-urokinase prodn. - using E.coli B strains and E.coli
 PT promoter P1rp and Shine-Dalgarno sequence MS-2.
 XX Disclosure; Page 7; 7pp; English.
 XX SER residue at position 21 is the start of the mature proUK.
 CC Non-glycosylated proUK (MW 45KD) produced by E.coli B strain containing
 CC the sequence.
 CC See also AAQ04101-07.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 431 AA;
 SQ Query Match 100.0%; Score 1508; DB 11; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPEELKFCQCGQKTLPRFKIIGGFTTIENQPFPAIYRRHRGGSVTVCGGSLIS 60
 DB 156 KPSSPEELKFCQCGQKTLPRFKIIGGFTTIENQPFPAIYRRHRGGSVTVCGGSLIS 215
 QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 275
 QY 121 IALLKIRKEGRCAPSRITQITCLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRKEGRCAPSRITQITCLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276
 DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 431
 RESULT 56
 AAR07112
 ID AAR07112 standard; protein; 431 AA.
 XX AAR07112;
 XX 25-MAR-2003 (updated)
 DT 24-JAN-1991 (first entry)
 XX Human pro-urokinase encoded by plasmid pUK1.
 XX pro-urokinase; transgenic mice.
 XX Synthetic.
 XX EP390592-A.
 XX 03-OCT-1990.
 XX 30-MAR-1990; 90EP-0303445.
 XX 31-MAR-1989; 89JP-0078574.
 XX (KYOW) KYOWA HAKKO KOGYO KK.

(EXPE-) CENT INST EXPR ANIMALS.
(JIKK-) JIKKEN DOBUTSU CHUO KENK.

PA Sekine S, Ito S, Katsuki M;
XX WPI; 1990-299492/40.
XX N-PSDB; AAQ06049.
XX Prodn. of recombinant protein, esp. human pro-urokinase - from
PT milk of transgenic animals using promoter of bovine alpha S1 casein
PT chromosomal gene.
XX Example; Table 1; 55pp; English.

XX E.coli strain C600Sf8 was transformed with recombinant plasmid
CC containing ds DNA derived from human pharynx cancer cell strain
CC Detroit 562. 10000 colonies were screened and one
CC positive clone was identified. Plasmid pUK1 was isolated and found
CC to contain the coding region and 3' non-coding region of pro-UK
CC downstream of Cys(41). Four silent substitutions were identified
CC (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC to
CC CCA; Gln(346), CAA to CAG.
CC See also AAQ06045-006048 and AAQ06392.
CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;
SQ Query Match 100.0%; Score 1508; DB 11; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 60
DB 156 KPSSPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 335
QY 181 VKLISHRECQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VKLISHRECQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 57
AAR47903
ID AAR47903 standard; Protein; 431 AA.
XX AC AAR47903;
XX 13-JUL-1994 (first entry)
XX DT Pro-urokinase derivative.
XX DE Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic;
XX KW factor.
XX OS Homo sapiens.
XX JP05336965-A.
XX 21-DEC-1993.
XX 17-OCT-1991; 91JP-0269615.

XX 17-OCT-1991; 91JP-0269615.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX WPI; 1994-030907/04.
XX N-PSDB; AAQ55772.
XX Novel human pro-urokinase deriva. having long half-life - with
PT high thrombolytic activity, useful for treatment of thrombosis
XX Disclosure; Page 15-17; 29pp; Japanese.
XX Sequences (AAQ55771-72) are pro-urokinase derivatives. The products
CC have an inserted sugar moiety having an amino acid substituted,
CC depleted or inserted variant around the thrombin cleavage site.
CC They also have a long half-life allowing them to be used in the
XX treatment of thrombosis.

SQ Sequence 431 AA;
Query Match 100.0%; Score 1508; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 60
DB 156 KPSSPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 335
QY 181 VKLISHRECQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VKLISHRECQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 58
AAR63141
ID AAR63141 standard; Protein; 431 AA.
XX AC AAR63141;
XX 25-MAR-2003 (updated)
XX DT 09-JUN-1995 (first entry)
XX DE Full length human urokinase protein.
XX KW Human urokinase glycoproteins; cardiovascular diseases;
XX pulmonary embolism.
XX OS Homo sapiens.
XX Key Location/Qualifiers
FT Sig peptide 1..20
FT Disulfide-bond 70..151
FT Disulfide-bond 91..133
FT Disulfide-bond 122..146
FT Disulfide-bond 168..299
FT Disulfide-bond 209..225
FT Disulfide-bond 217..288
FT Disulfide-bond 313..382
FT Disulfide-bond 345..361
FT Disulfide-bond 372..400

RESULT 60
 AAY50869
 ID AAY50869 standard; protein; 431 AA.
 XX AC AAY50869;
 XX DT 24-FEB-2000 (first entry)
 XX DE Human urokinase protein fragment.
 XX KW Urokinase; human; thrombolytic agent; streptokinase; antigenic;
 KW blood clot; heart attack; treatment.
 XX OS Homo sapiens.
 XX PN WO9957251-A2.
 XX PD 11-NOV-1999.
 XX PF 06-MAY-1999; 99WO-US10086.
 XX PR 06-MAY-1998; 98US-0084392.
 XX PA (OKLA-) OKLAHOMA MEDICAL RES FOUND.
 XX PI Zhang XC, Lin X, Tang JUN;
 XX DR WPI; 2000-052966/04.
 XX PT New thrombolytic agents derived from modified humanized streptokinase,
 PT useful for treating blood clot disorders -
 XX PS Disclosure; Page 46-48; 55pp; English.
 XX CC This invention describes a novel thrombolytic agent comprising
 CC streptokinase where at least one nonessential portion has been modified.
 CC The invention also describes a method of forming a thrombolytic agent
 CC comprising determining a nonessential portion of streptokinase and
 CC modifying the nonessential portion to render the resulting protein less
 CC antigenic. The modified streptokinase is used to treat blood clot
 CC disorders, such as heart attacks. The modified streptokinase has less
 CC antigenicity than streptokinase but is still able to complex plasminogen
 CC and lead to plasminogen activation. Modified streptokinase with the
 CC nonessential portions removed or truncated simplify the molecule. Such
 CC smaller proteins are cheaper and easier to produce. This sequence
 CC represents a fragment of the human urokinase protein which is used in
 CC the description of the method of the invention.
 XX SQ Sequence 431 AA;
 Query Match 100.0%; Score 1508; DB 21; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGKTLRPRFKIIGGFTTINQWPAAYRRHRGGSVTVCGSLIS 60
 DB 156 KPSSPPEELKFCQGKTLRPRFKIIGGFTTINQWPAAYRRHRGGSVTVCGSLIS 215
 QY 61 PCWVISATHCFIDYPKKEDIYVILGRSHLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDIYVILGRSHLNSNTQGMKFEVENLILHKDYSADTLAHND 275
 QY 121 IALLKIRKEGRCAOPSTIQTICLPVSYNDPFGTSCITGFGKENSTDYLYPEQLKMT 180
 DB 276 IALLKIRKEGRCAOPSTIQTICLPVSYNDPFGTSCITGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECOOPHYVIGSEVITKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTIGVS 240
 DB 336 VVKLISHRECOOPHYVIGSEVITKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTIGVS 395
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHYKEENGLAL 276

396 WGRGALKDKPGVYTRVSHFLPWIRSHYKEENGLAL 431

RESULT 61
 AAB84605
 ID AAB84605 standard; Protein; 431 AA.
 XX AC AAB84605;
 XX DT 05-SEP-2001 (first entry)
 XX DE Amino acid sequence of urokinase plasminogen activator.
 XX KW Growth factor; protein inhibitor; protease; damaged tissue;
 KW platelet-derived growth factor; PDGF; fibroblast growth factor; FGF;
 KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;
 KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
 KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
 KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
 KW vascular endothelial growth factor; urokinase plasminogen activator;
 KW dermal ulcer; wound.
 XX OS Homo sapiens.
 XX PN WO200149309-A2.
 XX PD 12-JUL-2001.
 XX PF 21-DEC-2000; 2000WO-IB01935.
 XX PR 29-DEC-1999; 99GB-0030768.
 XX PA (PFIZ) PFIZER LTD.
 XX PI (PFIZ) PFIZER INC.
 XX PI Davies MJ, Huggins JP, McIntosh FS, Occleston NL;
 XX WPI; 2001-418351/44.
 XX N-PSDB; AAH28220.
 XX CC Composition for the treatment of damaged tissue i.e. chronic wounds and
 CC dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
 CC factor -
 XX Disclosure; Page 550; 572pp; English.
 XX CC The specification describes a pharmaceutical composition, comprising
 CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
 CC agent inhibits the action of at least one specific adverse protein,
 CC i.e. a protease, that is upregulated in a damaged tissue such as a
 CC wound environment. Growth factors which are included in the composition
 CC of the invention are platelet-derived growth factor (PDGF), fibroblast
 CC growth factor (FGF), connective tissue derived growth factor (CTGF),
 CC keratinocyte-derived growth factor (KGF), transforming growth
 CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
 CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
 CC factor (VEGF), and chrysalin. Inhibitors which are included in the
 CC composition of the invention include inhibitors of urokinase-type
 CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
 CC composition is useful for the treatment of chronic damaged tissue, i.e.
 CC wounds and dermal ulcers. The present sequence represents a human uPA,
 CC and is used to produce the composition of the invention.
 XX SQ Sequence 431 AA;
 Query Match 100.0%; Score 1508; DB 22; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQGKTLRPRFKIIGGFTTINQWPAAYRRHRGGSVTVCGSLIS 60
 DB 156 KPSSPPEELKFCQGKTLRPRFKIIGGFTTINQWPAAYRRHRGGSVTVCGSLIS 215

PT Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
 PT useful for improving efficiency and reliability in drug development for
 PT treating thrombolytic disorders and cancer -
 PS Claim 27; Fig 3; 92pp; English.
 XX
 XX The invention relates to a polynucleotide comprising a first nucleotide
 CC sequence (NS1) comprising a PLAU (plasminogen activator, urokinase,
 CC a serine protease) isogene selected from isogenes 1-9 and 11-20 given
 CC in the specification, where each isogene comprises the regions of the
 CC PLAU gene or cDNA and is further defined by the corresponding sequence of
 CC polymorphisms (defining single nucleotide polymorphisms, SNP). Also
 CC included are methods of haplotyping/genotyping (and predicting the
 CC haplotype/genotype of the PLAU gene of an individual, identifying an
 CC association between a trait and at least one haplotype or haplotype pair
 CC of the PLAU gene, an isolated oligonucleotide for detecting a
 CC polymorphism in the PLAU gene, a recombinant non-human organism
 CC transformed or transfected with the gene or cDNA, fragments of the
 CC polynucleotides of at least 10 base pairs encompassing a polymorphic
 CC site, an isolated polymorphic variant PLAU protein or fragment, an
 CC isolated monoclonal antibody specific for PLAU, a computer system for
 CC storing and analysing polymorphism data for the PLAU gene and a genome
 CC anthology for the PLAU gene. PLAU is useful in screening for drugs
 CC targeting PLAU that are useful for treating thrombolytic disorders and
 CC cancers. The methods are useful for improving the efficiency and
 CC reliability of the discovery and development of drugs for treating
 CC diseases associated with PLAU activity, in validating PLAU as a drug
 CC target and in the design of clinical trials for treating a specific
 CC condition of disease associated with PLAU activity. The antibody is
 CC useful in diagnostic, prognostic and therapeutic methods. PLAU
 CC polynucleotides are useful in studying the expression and function of
 CC PLAU, and in expressing PLAU protein for use in screening for candidate
 CC drugs to treat diseases related to PLAU activity. The gene for PLAU
 CC is located on chromosome 10q24-qter. The present sequence represents the
 CC PLAU protein.
 XX
 XX Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSPPEELKFCQGOKTLPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60
 DB 156 KPSPPEELKFCQGOKTLPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 215
 QY 61 PCWVISATHCFIDYPKKEDIYVYVGRSLNSNTQGMKFEVENLILHKDYSDTLAHND 120
 DB 216 PCWVISATHCFIDYPKKEDIYVYVGRSLNSNTQGMKFEVENLILHKDYSDTLAHND 275
 QY 121 IALLKIRSKGRCAPSRITQICLPVNDPQFTSCITGFGKXENSTDVLYPEQLKMT 180
 DB 276 IALLKIRSKGRCAPSRITQICLPVNDPQFTSCITGFGKXENSTDVLYPEQLKMT 335
 QY 181 VVKLISHRECQPHYVGEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 240
 DB 336 VVKLISHRECQPHYVGEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGALKXDKPGVYTVRVSHFLPWIRSHTKENGLAL 276
 DB 396 WGRGALKXDKPGVYTVRVSHFLPWIRSHTKENGLAL 431

RESULT 64

ID AAU99232 standard; Protein; 431 AA.

AC AAU99232;

XX 24-SEP-2002 (first entry)

DE Human plasminogen activator, urokinase (PLAU) variant #4.

XX

KW Human: Plasminogen activator; urokinase; PLAU; cancer; enzyme;
 KW cystostatic; serine protease; thrombolytic disorder; isogene;
 KW pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;
 KW SNP; single nucleotide polymorphism; thrombolytic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 141
 FT /note= "Wild-type Pro substituted by Leu"
 XX
 XX WO200240503-A2.
 XX 23-MAY-2002.
 XX
 XX 14-NOV-2001; 2001WO-US44001.
 XX
 XX 17-NOV-2000; 2000US-249703P.
 XX (GENA-) GENAISSANCE PHARM INC.
 XX Anastasio AE, Bentivegna SC, Koshy B;
 XX WPI; 2002-519370/55.
 XX
 XX Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
 XX useful for improving efficiency and reliability in drug development for
 XX treating thrombolytic disorders and cancer -
 XX Claim 27; Page -; 92pp; English.

The invention relates to a polynucleotide comprising a first nucleotide
 sequence (NS1) comprising a PLAU (plasminogen activator, urokinase,
 a serine protease) isogene selected from isogenes 1-9 and 11-20 given
 in the specification, where each isogene comprises the regions of the
 PLAU gene or cDNA and is further defined by the corresponding sequence of
 polymorphisms (defining single nucleotide polymorphisms, SNP). Also
 included are methods of haplotyping/genotyping (and predicting the
 haplotype/genotype of the PLAU gene of an individual, identifying an
 association between a trait and at least one haplotype or haplotype pair
 of the PLAU gene, an isolated oligonucleotide for detecting a
 polymorphism in the PLAU gene, a recombinant non-human organism
 transformed or transfected with the gene or cDNA, fragments of the
 polynucleotides of at least 10 base pairs encompassing a polymorphic
 site, an isolated polymorphic variant PLAU protein or fragment, an
 isolated monoclonal antibody specific for PLAU, a computer system for
 storing and analysing polymorphism data for the PLAU gene and a genome
 anthology for the PLAU gene. PLAU is useful in screening for drugs
 targeting PLAU that are useful for treating thrombolytic disorders and
 cancers. The methods are useful for improving the efficiency and
 reliability of the discovery and development of drugs for treating
 diseases associated with PLAU activity, in validating PLAU as a drug
 target and in the design of clinical trials for treating a specific
 condition of disease associated with PLAU activity. The antibody is
 useful in diagnostic, prognostic and therapeutic methods. PLAU
 polynucleotides are useful in studying the expression and function of
 PLAU, and in expressing PLAU protein for use in screening for candidate
 drugs to treat diseases related to PLAU activity. The gene for PLAU
 is located on chromosome 10q24-qter. The present sequence
 represents a polymorphic variant of the PLAU protein.
 Note: The present sequence is not shown in the specification but
 was created by the indexer using the wild-type PLAU protein appearing as
 AAU99228 and the information on page 28.

Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSPPEELKFCQGOKTLPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60
 DB 156 KPSPPEELKFCQGOKTLPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 215

CC isolated monoclonal antibody specific for PLAU, a computer system for
CC storing and analysing polymorphism data for the PLAU gene and a genome
CC technology for the PLAU gene. PLAU is useful in screening for drugs
CC targeting PLAU that are useful for treating thrombotic disorders and
CC cancers. The methods are useful for improving the efficiency and
CC reliability of the discovery and development of drugs for treating
CC diseases associated with PLAU activity, in validating PLAU as a drug
CC target and in the design of clinical trials for treating a specific
CC condition of disease associated with PLAU activity. The antibody is
CC useful in diagnostic, prognostic and therapeutic methods. PLAU
CC polynucleotides are useful in studying the expression and function of
CC PLAU, and in expressing PLAU protein for use in screening for candidate
CC drugs to treat diseases related to PLAU activity. The gene for PLAU
CC is located on chromosome 10q24-qter. The present sequence
CC represents a polymorphic variant of the PLAU protein.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the wild-type PLAU protein appearing as
CC AAU99228 and the information on page 28.
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 23; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;

QY 1 KPSPPEBELKFCQCGKTLRPRFKIIGGEFTTIENQPMFAAIYRRHSGSVTVVCGSLIS 60
DB 156 KPSPPEBELKFCQCGKTLRPRFKIIGGEFTTIENQPMFAAIYRRHSGSVTVVCGSLIS 215

QY 61 PCWISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 335

QY 181 VKLISHRECQCPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VKLISHRECQCPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
DB 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 65
AAU99236
ID AAU99236 standard; Protein; 431 AA.
XX
AC AAU99236;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human plasminogen activator, urokinase (PLAU) variant #8.
XX
XX Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
XX cytosolic; serine protease; thrombotic disorder; isogene;
XX pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;
XX SNP; single nucleotide polymorphism; thrombotic; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 58
FT /note= "Wild-type Gly substituted by Arg"
XX
XX WO200240503-A2.
XX
PD 23-MAY-2002.
XX
XX

CC storing and analysing polymorphism data for the PLAU gene and a genome
CC technology for the PLAU gene. PLAU is useful in screening for drugs
CC targeting PLAU that are useful for treating thrombotic disorders and
CC cancers. The methods are useful for improving the efficiency and
CC reliability of the discovery and development of drugs for treating
CC diseases associated with PLAU activity, in validating PLAU as a drug
CC target and in the design of clinical trials for treating a specific
CC condition of disease associated with PLAU activity. The antibody is
CC useful in diagnostic, prognostic and therapeutic methods. PLAU
CC polynucleotides are useful in studying the expression and function of
CC PLAU, and in expressing PLAU protein for use in screening for candidate
CC drugs to treat diseases related to PLAU activity. The gene for PLAU
CC is located on chromosome 10q24-qter. The present sequence
CC represents a polymorphic variant of the PLAU protein.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the wild-type PLAU protein appearing as
CC AAU99228 and the information on page 28.
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 23; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;

QY 1 KPSPPEBELKFCQCGKTLRPRFKIIGGEFTTIENQPMFAAIYRRHSGSVTVVCGSLIS 60
DB 156 KPSPPEBELKFCQCGKTLRPRFKIIGGEFTTIENQPMFAAIYRRHSGSVTVVCGSLIS 215

QY 61 PCWISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWISATHCFIDYPKKEDYIVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMT 335

QY 181 VKLISHRECQCPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VKLISHRECQCPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276
DB 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 66
AAU99240
ID AAU99240 standard; Protein; 431 AA.
XX
AC AAU99240;
XX
DT 24-SEP-2002 (first entry)
XX
DE Human plasminogen activator, urokinase (PLAU) variant #12.
XX
XX Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
XX cytosolic; serine protease; thrombotic disorder; isogene;
XX pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;
XX SNP; single nucleotide polymorphism; thrombotic; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 58
FT /note= "Wild-type Gly substituted by Arg"
XX
XX Misc-difference 141
XX /note= "Wild-type Pro substituted by Leu"
XX
XX WO200240503-A2.
XX
PD 23-MAY-2002.
XX
XX

PF 14-NOV-2001; 2001WO-US44001.
 XX
 PR 17-NOV-2000; 2000US-249703P.
 XX
 PA (GENA-) GENAISSANCE PHARM INC.
 XX
 PI Anastasio AE, Bentivegna SC, Koshy B;
 XX
 DR WPI; 2002-519370/55.
 XX
 XX Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
 PT useful for improving efficiency and reliability in drug development for
 PT treating thrombolytic disorders and cancer -
 XX
 PS Claim 27; Page -; 92pp; English.
 XX
 CC The invention relates to a polynucleotide comprising a first nucleotide
 CC sequence (NSI) comprising a PLAU (plasminogen activator, urokinase,
 CC a serine protease) isogene selected from isogenes 1-9 and 11-20 given
 CC in the specification, where each isogene comprises the regions of the
 CC PLAU gene or cDNA and is further defined by the corresponding sequence of
 CC polymorphisms (defining single nucleotide polymorphisms, SNP). Also
 CC included are methods of haplotyping/genotyping (and predicting the
 CC haplotype/genotype of the PLAU gene of an individual, identifying an
 CC association between a trait and at least one haplotype or haplotype pair
 CC of the PLAU gene, an isolated oligonucleotide for detecting a
 CC polymorphism in the PLAU gene, a recombinant non-human organism
 CC transformed or transfected with the gene or cDNA, fragments of the
 CC polynucleotides of at least 10 base pairs encompassing a polymorphic
 CC site, an isolated polymorphic variant PLAU protein or fragment, an
 CC isolated monoclonal antibody specific for PLAU, a computer system for
 CC storing and analysing polymorphism data for the PLAU gene and a genome
 CC anthology for the PLAU gene. PLAU is useful in screening for drugs
 CC targeting PLAU that are useful for treating thrombolytic disorders and
 CC cancers. The methods are useful for improving the efficiency and
 CC reliability of the discovery and development of drugs for treating
 CC diseases associated with PLAU activity, in validating PLAU as a drug
 CC target and in the design of clinical trials for treating a specific
 CC condition of disease associated with PLAU activity. The antibody is
 CC useful in diagnostic, prognostic and therapeutic methods. PLAU
 CC polynucleotides are useful in studying the expression and function of
 CC PLAU, and in expressing PLAU protein for use in screening for candidate
 CC drugs to treat diseases related to PLAU activity. The gene for PLAU
 CC is located on chromosome 10q24-qter. The present sequence
 CC represents a polymorphic variant of the PLAU protein.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the wild-type PLAU protein appearing as
 CC AAU99228 and the information on page 28.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 1508; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVVCGSLIS 60
 Db 156 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVVCGSLIS 215
 QY 61 PCWVISATHCFIDYPPKEDYIVYIGRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120
 Db 216 PCWVISATHCFIDYPPKEDYIVYIGRSLNSNTQGMKFEVENLIHKDYSADTLAHND 275
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGKENSTDYLYPEQLKMT 180
 Db 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
 Db 336 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276

Db 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
 RESULT 67
 AAE17128
 ID AAE17128 standard; Protein; 431 AA.
 XX
 AC AAE17128;
 XX
 DT 18-APR-2002 (first entry)
 DE Human uPA protein.
 XX
 KW Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;
 KW Ets-1 transcription factor; N-acetylglucosaminyltransferase V; GNT-V;
 KW matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.
 XX
 OS Homo sapiens.
 OS
 PN WO200196606-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 14-JUN-2001; 2001WO-US19248.
 XX
 PR 14-JUN-2000; 2000US-0593488.
 XX
 PA (NYXI-) NYXIS NEURO THERAPIES INC.
 XX
 PI Yamamoto H, Kroes R, Moskal JR;
 DR WPI; 2002-130746/17.
 DR N-PSDB; AAD27855.
 XX
 PT Identifying a compound for treating cancer, comprises detecting
 PT transcription factor Ets-1, N-acetylglucosaminyltransferase V,
 PT urokinase-type plasminogen activator, matrix-type metalloproteinase-1
 PT and -3 gene expression -
 XX
 PS Example 1; Page 62-63; 63pp; English.
 XX
 CC The invention relates to a method of identifying a compound for treating
 CC cancer. The method involves detecting the expression of a panel of
 CC sequences selected from transcription factor Ets-1, urokinase-type
 CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),
 CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method
 CC is useful for identifying a compound that affects a cell, particularly a
 CC cancer cell or glioma cell, or a cell that is involved in inflammation.
 CC It is used for diagnosing and/or treating cancer or other conditions that
 CC are affected by one or more members of a panel of genes or their protein
 CC product. The method is also useful for drug discovery, drug safety
 CC evaluations and in gene therapy. The present sequence is human uPA
 CC protein.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 1508; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.4e-125;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVVCGSLIS 60
 Db 156 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVVCGSLIS 215
 QY 61 PCWVISATHCFIDYPPKEDYIVYIGRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120
 Db 216 PCWVISATHCFIDYPPKEDYIVYIGRSLNSNTQGMKFEVENLIHKDYSADTLAHND 275
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGKENSTDYLYPEQLKMT 180
 Db 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

||||| 336 VVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSGQSGPLVCSLQGRMTLTGIVS 395
Db
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 68
ABU56547
ID ABU56547 standard; Protein; 431 AA.
XX AC ABU56547;
XX 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #140.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US12476.
XX PR 18-APR-2001; 2001US-284770P.
XX PR 10-MAY-2001; 2001US-290492P.
XX PR 09-NOV-2001; 2001US-339245P.
XX PR 13-NOV-2001; 2001US-350666P.
XX PR 29-NOV-2001; 2001US-334370P.
XX PR 12-APR-2002; 2002US-372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX WPI; 2003-093161/08.
XX DR N-PSDB; ABX76275.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer -
XX PS Claim 27; Page 296; 453pp; English.
XX CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridises
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung
CC cancer-associated polynucleotides and polypeptides are used for
CC identifying a compound that modulates a lung cancer-associated
CC polypeptide, for inhibiting proliferation of a lung cancer-associated
CC cell to treat lung cancer in a patient and for treating a mammal having
CC lung cancer by administering a modulatory compound identified. The
CC methods are useful for treating lung cancer, such as small cell lung
CC cancer, non-small cell lung cancer or other benign or precancerous
CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
CC and polypeptides are useful for diagnostic purposes and as targets for
CC screening for therapeutic compounds that modulate lung cancer, such as
CC antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated
CC polypeptides of the invention.

SQ Sequence 431 AA;
Query Match 100.0%; Score 1508; DB 24; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPEELKFOGQKTLRPREKIIGGBFTTIENOPWFAAIYRRHRGGSVTVVCGSLIS 60
Db 156 KPSSPEELKFOGQKTLRPREKIIGGBFTTIENOPWFAAIYRRHRGGSVTVVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTTLAHND 120
Db 216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTTLAHND 275
QY 121 IALLKIRSKEGRCAPSRITQICLPYMYNDPQFTSCIEITGFGKENSTDYLYPEQLKMT 180
Db 276 IALLKIRSKEGRCAPSRITQICLPYMYNDPQFTSCIEITGFGKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSGQSGPLVCSLQGRMTLTGIVS 240
Db 336 VVKLISHRECQPHYYGSEVTTKMLCAADPWKTDSGQSGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431
RESULT 69
ABU56708
ID ABU56708 standard; Protein; 431 AA.
XX AC ABU56708;
XX DT 02-APR-2003 (first entry)
XX DE Lung cancer-associated polypeptide #301.
XX KW Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX OS Unidentified.
XX PN WO200286443-A2.
XX PD 31-OCT-2002.
XX PF 18-APR-2002; 2002WO-US12476.
XX PR 18-APR-2001; 2001US-284770P.
XX PR 10-MAY-2001; 2001US-290492P.
XX PR 09-NOV-2001; 2001US-339245P.
XX PR 13-NOV-2001; 2001US-350666P.
XX PR 29-NOV-2001; 2001US-334370P.
XX PR 12-APR-2002; 2002US-372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX WPI; 2003-093161/08.
XX DR N-PSDB; ABX76437.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer -
XX PS Claim 27; Page 424; 453pp; English.
XX CC The invention relates to a method for detecting a lung cancer-associated

transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the invention.

XX Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 24; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSPPEELKFCQCKQTLRPRFKIIGGFTTINQPFPAAIYRRHGGSVTVVCGSLIS 60
DB 156 KPSPPEELKFCQCKQTLRPRFKIIGGFTTINQPFPAAIYRRHGGSVTVVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
QY 121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 70

ABU11076
ID ABU11076 standard; Protein; 431 AA.

XX AC ABU11076;

XX DT 05-FEB-2003 (first entry)

XX DE Human urokinase plasminogen activator.

XX KW Urokinase plasminogen activator; gene therapy; cancer;
KW hyperproliferative disorder; cancer; breast cancer; colon cancer;
KW bone cancer; brain cancer; ovary cancer; cervix cancer;
KW endometrium cancer; stomach cancer; kidney cancer; tumor metastasis.

XX OS Homo sapiens.

XX PN WO200279515-A1.

XX PD 10-OCT-2002.

XX PF 18-MAR-2002; 2002WO-US08112.

XX PR 30-MAR-2001; 2001US-0821972.

XX PA (ISIS-) ISIS PHARM INC.

XX

PI Baker BP, Freier SM, Watt AT;
XX
DR WPI; 2003-058441/05.
DR N-PSDS; ABX17681.

XX New antisense compound, useful for preparing a composition for treating hyperproliferative disorders, cancer e.g., breast, colon, bone, brain, ovary, cervix, endometrium, stomach or kidney cancer, or tumor metastasis -

PS Disclosure; Page 101-102; 159pp; English.

XX A new compound, which is 8-50 nucleobases in length targeted to a nucleic acid molecule encoding urokinase plasminogen activator, specifically hybridises with and inhibits the expression of urokinase plasminogen activator. The compound is useful for preparing a composition for treating (e.g. by gene therapy) hyperproliferative disorder, cancer e.g., breast, colon, bone, brain, ovary, cervix, endometrium, stomach or kidney cancer, or tumor metastasis. This is the amino acid sequence of a urokinase plasminogen activator.

SQ Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 24; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSPPEELKFCQCKQTLRPRFKIIGGFTTINQPFPAAIYRRHGGSVTVVCGSLIS 60
DB 156 KPSPPEELKFCQCKQTLRPRFKIIGGFTTINQPFPAAIYRRHGGSVTVVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
QY 121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 71

AAR99888
ID AAR99888 standard; peptide; 432 AA.

XX AC AAR99888;

XX DT 27-JAN-1997 (first entry)

XX DE M43: fibrinolytic and anticoagulant activity contg. protein.

XX KW Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen;
KW urokinase; activator; streptokinase; staphylokinase; APSAC;
KW anisolated plasminogen streptokinase activator complex; hirudin;
KW hirullin; antistatin; pWLT27; pWS1; pSE8; pWS56.

XX OS Synthetic.

XX PN EP712934-A2.

XX PD 22-MAY-1996.

XX PF 03-NOV-1995; 95EP-0117316.

XX PR 17-NOV-1994; 94DE-4440892.

XX

PA (CHEF) GRUENENTHAL GMBH.
XX
XX PI Heinzl-Wieland R, Steffens GJ, Wnendt S;
XX DR WPI; 1996-240720/25.
XX
XX Proteins with fibrinolytic and anticoagulant activity - useful as
FT thrombolytic agents
FT
XX
XX PS Disclosure; Fig 21; 59pp; German.
XX
XX New peptides (I) with fibrinolytic and anticoagulant activity
CC comprise a plasminogen-activating amino acid sequence (A) fused
CC at the N- and/or C-terminus to a thrombin and/or factor Xa
CC inhibiting amino acid sequence (B). Excluded from the claims
CC are (i) where (A) is Ser47 to Leu411 of unglycosylated urokinase
CC linked at the C-terminus to sequences (i) to (iii):
CC T1-RP-T2-GGGGNGDFEIPBEYL-T3
CC T1-RPFLNPNNDKVPFWEDEKNE (ii)
CC T1-RPSSEFEPEIDEEX (iii)
CC Where T1= P or V; T2 = L or a bond; T3= Q or OH.
CC (A) is pref. (pro)urokinase; tissue plasminogen activator (tPA),
CC bat-PA (all opt. modified by deletion, substitution, insertion and/or
CC adn.); streptokinase; staphylokinase; and/or APSAC (anisolated
CC plasminogen streptokinase activator complex), esp. prourokinase
CC (411 amino acids) or its Ser47 to Leu411 or Ser138 to Leu411
CC fragments, or t-PA (527 amino acids) or its Ser89Arg to 527Pro or
CC 174Ser to 527Pro fragments.
CC (B) has hirudin or hirullin activities; or is derived from the human
CC thrombin receptor, antistatin and/or the tick anticoagulant peptide.
CC Most pref. are the 65 amino acid hirudin sequence or one of the six
CC sequences given in AAR99879 to AAR99884.
CC Plasminids pMT27 (M51), pWS1 (M5112), pSE8 (M36) and pHW56 (M43)
CC contain the sequences encoding AAR99885 to AAR99888, respectively.
CC The products were tested in human citrated plasma (5 microg in 200
CC microl 1:10 diluted plasma). The thrombin time was then 1.2, 3, 2.8
CC and 1.2 times greater, respectively, than in the absence of the product.
XX
XX SQ Sequence 432 AA;
Query Match 100.0%; Score 1508; DB 17; Length 432;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVYVCGSLIS 60
DB 157 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVYVCGSLIS 216
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120
DB 217 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 276
QY 121 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 180
DB 277 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 336
QY 181 VVKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 337 VVKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 396
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 276
DB 397 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 432
RESULT 72
AAR20537
ID AAR20537 standard; Protein; 434 AA.
XX
XX AC AAR20537;
XX
XX DT 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)

XX Amidated deriv. of pro-urokinase (2).
DE
XX Pro-urokinase; plasminogen activator; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT Protein 1..431
FT /label= pro-UK
XX
XX DE4122688-A.
PN
XX 16-JAN-1992.
PD
XX 09-JUL-1991; 91DE-4122688.
PF
XX 12-JUL-1990; 90GB-0015369.
PR 10-JUL-1991; 91GB-0014846.
XX
XX (FARM) FARMITALIA ERBA SRL CARLO.
PA
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
XX WPI; 1992-025815/04.
PI
XX New amidated derivs. of human pro-urokinase - are fibrinolytic
PT and can be used to treat acute myocardial infarction, pulmonary
PT embolism or deep venous thrombosis
XX
XX Claim 4,8; Page 8; 18pp; German.
PS
XX The protein has fibrinolytic activity and can be used in the same
CC way as PUK, e.g. for treating acute myocardial infarction, lung
CC embolism and deep venous thrombosis. It has greater affinity for
CC plasminogen bound to fibrin than for circulating plasminogen, so
CC have high selectivity for thrombi with reduced chance of bleeding.
CC Compared with the COOH-terminated cpds., it has a better stability
CC against most carboxypeptidases and prolonged half life.
CC See also AAQ20360, AAQ20754-58 and AAR20536-38.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 434 AA;
Query Match 100.0%; Score 1508; DB 13; Length 434;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVYVCGSLIS 60
DB 156 KPSSPPEELKFQCGQKTLRPFKIIIGGEFTTIENQPFALYRRHGGSVTVYVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHND 275
QY 121 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECCQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 276
DB 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 431
RESULT 73
AAR20538
ID AAR20538 standard; Protein; 434 AA.
XX

```

AC AAR20538;
XX
XX 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)
XX
XX Amidated deriv. of pro-urokinase (3).
DE
XX Pro-urokinase; plasminogen activator; ss.
KW
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..431
FT /label= pro-UK
FT
XX DE4122688-A.
XX
XX 16-JAN-1992.
XX
XX 09-JUL-1991; 91DE-4122688.
XX
XX 12-JUL-1990; 90GB-0015369.
XX 10-JUL-1991; 91GB-0014846.
XX (PAM ) FARMITALIA ERBA SRL CARLO.
XX
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
XX WPI; 1992-025815/04.
XX
XX New amidated derivs. of human pro-urokinase - are fibrinolytic
XX and can be used to treat acute myocardial infarction, pulmonary
XX embolism or deep venous thrombosis
XX
XX Claim 4, 9; Page 8; 18pp; German.
XX
XX The protein has fibrinolytic activity and can be used in the same
XX way as PUK, e.g. for treating acute myocardial infarction, lung
XX embolism and deep venous thrombosis. It has greater affinity for
XX plasminogen bound to fibrin than for circulating plasminogen, so
XX have high selectivity for thrombi with reduced chance of bleeding.
XX Compared with the COOH-terminated cpds., it has a better stability
XX against most carboxypeptidases and prolonged half life.
XX See also AAQ20360, AAQ20754-58 and AAR20536-38.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 434 AA;

Query Match 100.0%; Score 1508; DB 13; Length 434;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSPPPEELKFCQCKTLRPRFKIIGGFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60
DB 156 KPSPPPEELKFCQCKTLRPRFKIIGGFTTIENQPFPAIYRRHGGSVTVVCGSLIS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSTDYLYPEQLKMT 335

QY 181 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKTDCQGDGSGPLVCSLQGRMTLGIYS 240
DB 336 VVKLISHRECCQPHYGYSEVTTKMLCAADPQWKTDCQGDGSGPLVCSLQGRMTLGIYS 395

QY 241 WGRGCALKDQPGYVTVRSHFLPWIRSHTEKENGIAL 276
DB 396 WGRGCALKDQPGYVTVRSHFLPWIRSHTEKENGIAL 431

```

```

RESULT 74
AAR20536
ID AAR20536 standard; Protein; 436 AA.
XX
XX AAR20536;
XX
XX 25-MAR-2003 (updated)
DT 21-MAY-1992 (first entry)
XX
XX Amidated deriv. of pro-urokinase (1).
DE
XX Pro-urokinase; plasminogen activator; ss.
KW
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..431
FT /label= pro-UK
FT /note= "the amidated form of pro-UK is
FT claimed in claim 1"
FT Misc-difference 433..436
FT /note= "may be any amino acid, pref. Lys or Arg,
FT or 0-4 amino acids may be omitted"
XX
XX DE4122688-A.
XX
XX 16-JAN-1992.
XX
XX 09-JUL-1991; 91DE-4122688.
XX
XX 12-JUL-1990; 90GB-0015369.
XX 10-JUL-1991; 91GB-0014846.
XX (PAM ) FARMITALIA ERBA SRL CARLO.
XX
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
XX WPI; 1992-025815/04.
XX
XX New amidated derivs. of human pro-urokinase - are fibrinolytic
XX and can be used to treat acute myocardial infarction, pulmonary
XX embolism or deep venous thrombosis
XX
XX Claim 1, 4, 7; Page 8; 18pp; German.
XX
XX The protein has fibrinolytic activity and can be used in the same
XX way as PUK, e.g. for treating acute myocardial infarction, lung
XX embolism and deep venous thrombosis. It has greater affinity for
XX plasminogen bound to fibrin than for circulating plasminogen, so
XX have high selectivity for thrombi with reduced chance of bleeding.
XX Compared with the COOH-terminated cpds., it has a better stability
XX against most carboxypeptidases and prolonged half life.
XX See also AAQ20360, AAQ20754-58 and AAR20536-38.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 436 AA;

Query Match 100.0%; Score 1508; DB 13; Length 436;
Best Local Similarity 100.0%; Pred. No. 2.4e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSPPPEELKFCQCKTLRPRFKIIGGFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60
DB 156 KPSPPPEELKFCQCKTLRPRFKIIGGFTTIENQPFPAIYRRHGGSVTVVCGSLIS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKENSTDYLYPEQLKMT 335

```

Qy	181	VVKLIHRECOOPHYGSEVTKMLCAADPOWKTDSCQDSGGPLVCSLQGRMTLTGIVS	240
Db	336	VVKLIHRECOOPHYGSEVTKMLCAADPOWKTDSCQDSGGPLVCSLQGRMTLTGIVS	395
Qy	241	WGRGCALDKPGVYTRVSHFLPWIRSHTKENGLAL	276
Db	396	WGRGCALDKPGVYTRVSHFLPWIRSHTKENGLAL	431

Search completed: January 8, 2004, 13:18:55
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2004, 12:58:25 ; Search time 32 Seconds
(without alignments)
1738.022 Million cell updates/sec

Title: US-09-880-503-5
Perfect score: 1508
Sequence: 1 KPSSPPEELKFCGQKTLRP.....VSHFLPWIRSHKEBGLAL 276

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747907 seqs, 201509753 residues

Total number of hits satisfying chosen parameters: 747907

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/FCIUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1508	100.0	276	10	US-09-880-503-5
2	1508	100.0	323	10	US-09-880-503-7
3	1508	100.0	411	10	US-09-880-503-3
4	1508	100.0	411	12	US-10-407-821-2
5	1508	100.0	431	12	US-10-301-822-161
6	1508	100.0	431	12	US-10-247-671-149
7	1508	100.0	431	12	US-10-131-995-21
8	1508	100.0	431	12	US-10-295-027-414
9	1508	100.0	431	12	US-10-295-027-1275
10	1508	100.0	431	14	US-10-076-421-2
11	1508	100.0	431	15	US-10-171-311-184
12	1505	98.8	431	15	US-10-193-856-4
13	1490	98.8	431	10	US-09-264-468B-1
14	1465	97.1	403	10	US-09-880-503-6
15	1464	97.1	268	12	US-10-407-821-3

16	1333	88.4	246	10	US-09-264-468B-2	Sequence 2, Appli
17	1318	87.4	241	11	US-09-898-837A-47	Sequence 47, Appl
18	742	49.2	337	15	US-10-106-698-6266	Sequence 6266, Ap
19	591	39.2	118	12	US-09-997-003-39	Sequence 39, Appl
20	583	38.7	268	11	US-09-987-457-17	Sequence 17, Appl
21	583	38.7	268	11	US-09-987-455-18	Sequence 18, Appl
22	583	38.7	308	11	US-09-987-457-16	Sequence 16, Appl
23	583	38.7	308	11	US-09-987-455-17	Sequence 17, Appl
24	583	38.7	331	11	US-09-987-457-11	Sequence 11, Appl
25	583	38.7	331	11	US-09-987-455-12	Sequence 12, Appl
26	583	38.7	335	11	US-09-987-457-13	Sequence 13, Appl
27	583	38.7	335	11	US-09-987-455-14	Sequence 14, Appl
28	583	38.7	339	11	US-09-987-457-12	Sequence 12, Appl
29	583	38.7	339	11	US-09-987-455-13	Sequence 13, Appl
30	583	38.7	343	11	US-09-987-457-14	Sequence 14, Appl
31	583	38.7	343	11	US-09-987-455-15	Sequence 15, Appl
32	583	38.7	343	11	US-09-987-457-15	Sequence 15, Appl
33	583	38.7	343	11	US-09-987-455-16	Sequence 16, Appl
34	583	38.7	354	11	US-09-987-457-10	Sequence 10, Appl
35	583	38.7	354	11	US-09-987-455-11	Sequence 11, Appl
36	583	38.7	372	9	US-09-084-491A-3	Sequence 3, Appli
37	583	38.7	372	14	US-10-102-704-3	Sequence 3, Appli
38	583	38.7	377	11	US-09-987-455-8	Sequence 8, Appli
39	583	38.7	527	11	US-09-987-457-18	Sequence 18, Appl
40	583	38.7	527	11	US-09-987-455-19	Sequence 19, Appl
41	583	38.7	562	9	US-09-969-271-7	Sequence 7, Appli
42	583	38.7	562	10	US-09-974-298-145	Sequence 145, App
43	583	38.7	562	12	US-10-443-701-4	Sequence 4, Appli
44	583	38.7	562	15	US-10-193-656-8	Sequence 8, Appli
45	583	38.7	650	12	US-10-401-077-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-880-503-5
; Sequence 5, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-5

Query Match	100.0%;	Score	1508;	DB	10;	Length	276;
Best Local Similarity	100.0%;	Pred. No.	4.2e-146;				
Matches	276;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	KPSSPPEELKFCGQKTLRPFKIIGGEFTTIENQWFAAIYRRHGGSVYVCGGSLIS	60				
Db	1	KPSSPPEELKFCGQKTLRPFKIIGGEFTTIENQWFAAIYRRHGGSVYVCGGSLIS	60				
Qy	61	PCWISATHCFIDYPPKEDYIVLGRSLNSNTQGMKFEVNLILHKYSADTLAHND	120				
Db	61	PCWISATHCFIDYPPKEDYIVLGRSLNSNTQGMKFEVNLILHKYSADTLAHND	120				
Qy	121	IALLKIRSEKRCQAQPSRTIQTICLPSWMYNDPQGTSCETIGFGKENSTDYLYPEQLKMT	180				
Db	121	IALLKIRSEKRCQAQPSRTIQTICLPSWMYNDPQGTSCETIGFGKENSTDYLYPEQLKMT	180				

Qy 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

RESULT 2
US-09-880-503-7
; Sequence 7, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-7

Query Match 100.0%; Score 1508; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 5.2e-146;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFQCGQKTLRPRFKIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLIS 60
Db 48 KPSSPPEELKFQCGQKTLRPRFKIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLIS 107
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYSADTLAHND 120
Db 108 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYSADTLAHND 167
Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 180
Db 168 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 227
Qy 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 228 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 287
Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 288 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 323

RESULT 3
US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3

; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 100.0%; Score 1508; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 7.2e-146;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFQCGQKTLRPRFKIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLIS 60
Db 136 KPSSPPEELKFQCGQKTLRPRFKIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLIS 195
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYSADTLAHND 120
Db 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYSADTLAHND 255
Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 180
Db 256 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 315
Qy 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 316 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 376 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 4
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 100.0%; Score 1508; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 7.2e-146;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFQCGQKTLRPRFKIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLIS 60
Db 136 KPSSPPEELKFQCGQKTLRPRFKIIGGETTIENQWPFAAIYRRHRGGSVTVVCGSLIS 195
Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYSADTLAHND 120
Db 196 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYSADTLAHND 255
Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 180
Db 256 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENSTDYLYPEQLKMT 315
Qy 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 316 VKLISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 5

US-10-301-822-161
 ; Sequence 161, Application US/10301822
 ; Publication No. US20030148410A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Berger, Allison
 ; APPLICANT: Guillemette, Tracy L.
 ; APPLICANT: Kamackar, Shubhangi
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Thibodeau, Stephen N.
 ; APPLICANT: Burgart, Lawrence J.
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; THERAPY OF COLON CANCER
 ; FILE REFERENCE: MP001-029P2RNM
 ; CURRENT APPLICATION NUMBER: US/10/301,822
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 60/339,971
 ; PRIOR FILING DATE: 2001-12-10
 ; PRIOR APPLICATION NUMBER: US 60/361,978
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/381,988
 ; PRIOR FILING DATE: 2002-05-20
 ; NUMBER OF SEQ ID NOS: 228
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 161
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-10-301-822-161

Query Match 100.0%; Score 1508; DB 12; Length 431;
 Best Local Similarity 100.0%; Pred. No. 7,6e-146;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRHRGGSVTVVCGSLIS 60
 Db 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRHRGGSVTVVCGSLIS 215

QY 61 PCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 216 PCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180
 Db 276 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 335

QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKTDCQGDSCGGLVCSLQGRMTLTGIVS 240
 Db 336 VVKLISHRECQPHYGVSVTTKMLCAADPQWKTDCQGDSCGGLVCSLQGRMTLTGIVS 395

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6

US-10-247-671-149
 ; Sequence 149, Application US/10247671
 ; Publication No. US20030194721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mkita, Thomas
 ; APPLICANT: Shiffman, Dov
 ; APPLICANT: Porter, Gordon, J.
 ; APPLICANT: Kaser, Matthew R.

; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
 ; FILE REFERENCE: PA-0050 US
 ; CURRENT APPLICATION NUMBER: US/10/247,671
 ; CURRENT FILING DATE: 2002-09-18
 ; PRIOR APPLICATION NUMBER: 60/323,784
 ; PRIOR FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 149
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
 US-10-247-671-149

Query Match 100.0%; Score 1508; DB 12; Length 431;
 Best Local Similarity 100.0%; Pred. No. 7,6e-146;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRHRGGSVTVVCGSLIS 60
 Db 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQWPFAAIYRHRGGSVTVVCGSLIS 215

QY 61 PCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 216 PCWISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180
 Db 276 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 335

QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKTDCQGDSCGGLVCSLQGRMTLTGIVS 240
 Db 336 VVKLISHRECQPHYGVSVTTKMLCAADPQWKTDCQGDSCGGLVCSLQGRMTLTGIVS 395

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 7

US-10-131-985-21
 ; Sequence 21, Application US/10131985
 ; Publication No. US2003019940A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dack, Kevin N
 ; APPLICANT: Davies, Michael J
 ; APPLICANT: Fish, Paul V
 ; APPLICANT: Huggins, Jonathan P
 ; APPLICANT: McIntosh, Fraser S
 ; APPLICANT: Occleston, Nicholas L
 ; TITLE OF INVENTION: Composition
 ; FILE REFERENCE: PCS 10391A
 ; CURRENT APPLICATION NUMBER: US/10/131,985
 ; CURRENT FILING DATE: 2002-04-25
 ; PRIOR APPLICATION NUMBER: US/09/726,295
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR APPLICATION NUMBER: GB 9930768.8
 ; PRIOR FILING DATE: 1999-12-29
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-131-985-21

Query Match 100.0%; Score 1508; DB 12; Length 431;
 Best Local Similarity 100.0%; Pred. No. 7,6e-146;
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQKTLRPRFKIIGBFTTIENQWFAAIYRRHRGGSVTVYCGGSLIS 60
DB 156 KPSSPPEELKFCQCKQKTLRPRFKIIGBFTTIENQWFAAIYRRHRGGSVTVYCGGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120
DB 216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 275
QY 121 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 276
DB 396 WGRGCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 431

RESULT 8

US-10-295-027-414
; Sequence 414, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 414
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-414

Query Match 100.0%; Score 1508; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.6e-146;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKQKTLRPRFKIIGBFTTIENQWFAAIYRRHRGGSVTVYCGGSLIS 60
DB 156 KPSSPPEELKFCQCKQKTLRPRFKIIGBFTTIENQWFAAIYRRHRGGSVTVYCGGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120
DB 216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 275
QY 121 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSEKRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGFKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 276
DB 396 WGRGCALKDKPGVYTVRVSHFLPWIRSHTKENGLAL 431

RESULT 9

US-10-295-027-1275
; Sequence 1275, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1275
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1275

Query Match 100.0%; Score 1508; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.6e-146;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFWAAIYRRHRGGSVTVVCGSGLIS 60
Db 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFWAAIYRRHRGGSVTVVCGSGLIS 215
Qy 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHND 275
Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 335
Qy 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLOGRMTLTGIVS 240
Db 336 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLOGRMTLTGIVS 395
Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 10
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match 100.0%; Score 1508; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.6e-146;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFWAAIYRRHRGGSVTVVCGSGLIS 60
Db 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFWAAIYRRHRGGSVTVVCGSGLIS 215
Qy 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHND 275
Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 335
Qy 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLOGRMTLTGIVS 240
Db 336 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLOGRMTLTGIVS 395
Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 11
US-10-171-311-184

; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match 100.0%; Score 1508; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.6e-146;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFWAAIYRRHRGGSVTVVCGSGLIS 60
Db 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFWAAIYRRHRGGSVTVVCGSGLIS 215
Qy 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHND 120
Db 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHND 275
Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180
Db 276 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFKENSTDYLYPEQLKMT 335
Qy 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLOGRMTLTGIVS 240
Db 336 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLOGRMTLTGIVS 395
Qy 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 12
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/13577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182

; PRIOR FILING DATE: 2001-07-13
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: GenBank / P00749
 ; DATABASE ENTRY DATE: 1986-07-21
 ; RELEVANT RESIDUES: (1)..(431)
 ; US-10-193-656-4

Query Match 99.8%; Score 1505; DB 15; Length 431;
 Best Local Similarity 99.6%; Pred. No. 1.6e-145;
 Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVYCGGSLIS 60
 Db 156 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVYCGGSLIS 215
 QY 61 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 216 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
 QY 121 IALLKIRSKGRCQAPSRITQICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
 Db 276 IALLKIRSKGRCQAPSRITQICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDCQGDGGPLVCSLQGRMTLTGIVS 240
 Db 336 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDCQGDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 396 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 13
 US-09-264-468B-1
 ; Sequence 1, Application US/09264468B
 ; Patent No. US20020106775A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Jieyi
 ; APPLICANT: Nienaber, Vicki L.
 ; APPLICANT: Henkin, Jack
 ; APPLICANT: Smith, Richard A.
 ; APPLICANT: Walter, Karl A.
 ; APPLICANT: Severin, Jean M.
 ; APPLICANT: Edalji, Rohinton
 ; APPLICANT: Johnson Jr., Robert W.
 ; APPLICANT: Holzman, Thomas F.
 ; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
 ; FILE REFERENCE: 6310-US.PI
 ; CURRENT APPLICATION NUMBER: US/09/264,468B
 ; CURRENT FILING DATE: 1999-03-05
 ; PRIOR APPLICATION NUMBER: US 09/036,361
 ; PRIOR FILING DATE: 1998-03-06
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 431
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SIGNAL
 ; LOCATION: (1)..(20)
 ; OTHER INFORMATION: Leader sequence
 ; NAME/KEY: VARIANT
 ; LOCATION: (279)..(279)
 ; OTHER INFORMATION: Xaa = any amino acid
 ; NAME/KEY: VARIANT
 ; LOCATION: (302)..(302)

; OTHER INFORMATION: Xaa = any amino acid
 ; US-09-264-468B-1

Query Match 98.8%; Score 1490; DB 10; Length 431;
 Best Local Similarity 99.3%; Pred. No. 5.3e-144;
 Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVYCGGSLIS 60
 Db 156 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVYCGGSLIS 215
 QY 61 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
 Db 216 PCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
 QY 121 IALLKIRSKGRCQAPSRITQICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
 Db 276 IALLKIRSKGRCQAPSRITQICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 335
 QY 181 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDCQGDGGPLVCSLQGRMTLTGIVS 240
 Db 336 VVKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDCQGDGGPLVCSLQGRMTLTGIVS 395
 QY 241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 396 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 14
 US-09-880-503-6
 ; Sequence 6, Application US/09880503
 ; Patent No. US20020131964A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CINES, Douglas B
 ; APPLICANT: HIGAZI, Abd Al-Roof
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
 ; TITLE OF INVENTION: TISSUE CONTRACTILITY
 ; FILE REFERENCE: 9596-331
 ; CURRENT APPLICATION NUMBER: US/09/880,503
 ; CURRENT FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/212,847
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 403
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-880-503-6

Query Match 97.1%; Score 1465; DB 10; Length 403;
 Best Local Similarity 99.6%; Pred. No. 1.8e-141;
 Matches 268; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 8 ELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVYCGGSLISPCWVISA 67
 Db 135 ELKFCQCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVYCGGSLISPCWVISA 194
 QY 68 THCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIR 127
 Db 195 THCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIR 254
 QY 128 SKGRCQAPSRITQICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMTVVKLISH 187
 Db 255 SKGRCQAPSRITQICLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMTVVKLISH 314
 QY 188 RECQOPHYGVSEVTTKMLCAADPOWKTDCQGDGGPLVCSLQGRMTLTGIVSWGRCAL 247
 Db 315 RECQOPHYGVSEVTTKMLCAADPOWKTDCQGDGGPLVCSLQGRMTLTGIVSWGRCAL 374
 QY 248 KDKPGVYTRVSHFLPWIRSHTKENGLAL 276
 Db 375 KDKPGVYTRVSHFLPWIRSHTKENGLAL 403

RESULT 15
US-10-407-821-3
; Sequence 3, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPELURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; FILE REFERENCE: UTSN:02205
; CURRENT FILING DATE: 2003-04-04
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-3

Query Match 97.1%; Score 1464; DB 12; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.3e-141;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 LKFOGQKTLRPFKLIIGFEFTIENQPFALYRRHGGSVTYVCGSLISPCWVISAT 68
Db 1 LKFOGQKTLRPFKLIIGFEFTIENQPFALYRRHGGSVTYVCGSLISPCWVISAT 60
QY 69 HCFIDYPKKEDIYVILGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRS 128
Db 61 HCFIDYPKKEDIYVILGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRS 120
QY 129 KEGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHR 188
Db 121 KEGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHR 180
QY 189 EQQPHYGVSEVTKMLCAADPQWKTDCQSGSGPLVCSLQGRMTLTGIVSWGRGCALK 248
Db 181 EQQPHYGVSEVTKMLCAADPQWKTDCQSGSGPLVCSLQGRMTLTGIVSWGRGCALK 240
QY 249 DRPGVYTRVSHFLPWRSHRSTKEENGLAL 276
Db 241 DRPGVYTRVSHFLPWRSHRSTKEENGLAL 268

RESULT 16
US-09-264-468B-2
; Sequence 2, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nieraber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-468B-2
Query Match 88.4%; Score 1333; DB 10; Length 246;
Best Local Similarity 99.2%; Pred. No. 3.1e-138;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 24 IIGFEFTIENQPFALYRRHGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVY 83
Db 1 IIGFEFTIENQPFALYRRHGGSVTYVCGSLISPCWVISATHCFIDYPKKEDIYVY 60
QY 84 LGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIQT 143
Db 61 LGRSLNSNTGEMKFEVENILHKDYSADTLAHNDIALLKIRSKGRCACQPSRTIQT 120
QY 144 CLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTK 203
Db 121 ALPSMYNDPQGTSCETITGFGKEQSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTK 180
QY 204 MLCADPQWKTDCQSGSGPLVCSLQGRMTLTGIVSWGRGCALKDRPGVYTRVSHFLPW 263
Db 181 MLCADPQWKTDCQSGSGPLVCSLQGRMTLTGIVSWGRGCALKDRPGVYTRVSHFLPW 240
QY 264 IRSHTK 269
Db 241 IRSHTK 246

RESULT 17
US-09-898-837A-47
; Sequence 47, Application US/09898837A
; Publication No. US20030077697A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Herrmann, John L.
; APPLICANT: Burgess, Catherine
; APPLICANT: Fernandes, Elma
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Rastelli, Luca
; APPLICANT: Curagen Corporation
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: MacDougall, John R.
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: 15966-598 CIP
; CURRENT APPLICATION NUMBER: US/09/898,837A
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 47
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-898-837A-47

Query Match      87.4%; Score 1318; DB 11; Length 241;
Best Local Similarity 99.6%; Pred. No. 1e-126;
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 24 IIGGEFTTIENQPFALYRHRGGSVTVYCGSLISPCWVISATHCFIDYPKKEDVIVY 83
DB 1 IIGGEFTTIENQPFALYRHRGGSVTVYCGSLMSPCWVISATHCFIDYPKKEDVIVY 60

QY 84 LGRSLNSNTQGMKFVENILHKDYSADTLAHNDIALKIRSKRGCAQPSRTTQTI 143
DB 61 LGRSLNSNTQGMKFVENILHKDYSADTLAHNDIALKIRSKRGCAQPSRTTQTI 120

QY 144 CLPSMYNDPQGTSCETGFCNKSTVDLYPEQLKMTVVKLISHRECQOPHYGSEVTTK 203
DB 121 CLPSMYNDPQGTSCETGFCNKSTVDLYPEQLKMTVVKLISHRECQOPHYGSEVTTK 180

QY 204 MLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPQVYTRVSHFLPW 263
DB 181 MLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPQVYTRVSHFLPW 240

QY 264 I 264
DB 241 I 241

RESULT 18
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match      49.2%; Score 742; DB 15; Length 337;
Best Local Similarity 99.9%; Pred. No. 1.7e-67;
Matches 140; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 1 KPSSPEELKFCGOKTLRPRFKIIGGFTTIENQPFALYRHRGGSVTVYCGSLIS 60
DB 162 KPSSPEELKFCGOKTLRPRFKIIGGFTTIENQPFALYRHRGGSVTVYCGSLIS 221

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENILHKDYSADTLAHND 120
DB 222 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENILHKDYSADTLAHND 281

QY 121 IALLKIRSKRGCAQ-----PSRTTIQICPSM 148
DB 282 IALLKIRSKRGCAQHPLGRYSACPRCTITPSL 315

RESULT 19
```

```
US-09-997-003-39
; Sequence 39, Application US/09997003
; Publication No. US20030203361A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA003PI
; CURRENT APPLICATION NUMBER: US/09/997,003
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/22157
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/148,680
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 39
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-003-39

Query Match      39.2%; Score 591; DB 12; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.3e-52;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 DYLPQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSL 229
DB 12 DYLPQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSL 71

QY 230 QGRMTLTGIVSWGRCALKDKPQVYTRVSHFLPWIRSHTKENGLAL 276
DB 72 QGRMTLTGIVSWGRCALKDKPQVYTRVSHFLPWIRSHTKENGLAL 118

RESULT 20
US-09-987-457-17
; Sequence 17, Application US/09987457
; Publication No. US20030013150A1
; ( NEURAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 17
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2s molecule (modified)
US-09-987-457-17

Query Match      38.7%; Score 583; DB 11; Length 268;
Best Local Similarity 44.9%; Pred. No. 2.5e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFALYRHRH-RGGSVTVYCGSLISPCWVISATHC 70
DB 5 CGLRQYSQPOFRIKGGLFADIASHPQQAIFAKHRRSPGRFLCGGILISSCWILSAAC 64
```

```
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILILHKDYSADTLAHNDIALLKIRSK 130
Db 65 FQERPPHLLTVILGRTYRVVPGEEQKEVEKYIVHKEFDDT--YDNDIALQLKSDS 122
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSBITGKESNTDYLYPEQLKMTVVKLISHREC 190
Db 123 SRCQAESVVRTVCLPPADLQPLDPTWTECELSGYGKEALSPFYSERLKEAHRVLPSSRC 182
QY 191 QPPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLOGRMTLTGIVSWG 244
Db 183 TSOHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLISWGLG 242
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 243 CGQKDPGVYTKVTNYLDWIRDNR 267

RESULT 21
US-09-987-455-18
; Sequence 18, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tavapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; K2S molecule
US-09-987-455-18

Query Match 38.7%; Score 583; DB 11; Length 268;
Best Local Similarity 44.9%; Pred. No. 2.5e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPFKIIIGGFTTIENQPMFAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
Db 5 CGLRQYSQPFRIKGLGFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 64
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILILHKDYSADTLAHNDIALLKIRSK 130
Db 65 FQERPPHLLTVILGRTYRVVPGEEQKEVEKYIVHKEFDDT--YDNDIALQLKSDS 122
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSBITGKESNTDYLYPEQLKMTVVKLISHREC 190
Db 123 SRCQAESVVRTVCLPPADLQPLDPTWTECELSGYGKEALSPFYSERLKEAHRVLPSSRC 182
QY 191 QPPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLOGRMTLTGIVSWG 244
Db 183 TSOHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLISWGLG 242
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 243 CGQKDPGVYTKVTNYLDWIRDNR 267

RESULT 22
US-09-987-457-16
; Sequence 16, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tavapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 16
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; K2S molecule (modified)
US-09-987-457-16

Query Match 38.7%; Score 583; DB 11; Length 308;
Best Local Similarity 44.9%; Pred. No. 3e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPFKIIIGGFTTIENQPMFAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
Db 45 CGLRQYSQPFRIKGLGFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 104
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILILHKDYSADTLAHNDIALLKIRSK 130
Db 105 FQERPPHLLTVILGRTYRVVPGEEQKEVEKYIVHKEFDDT--YDNDIALQLKSDS 162
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSBITGKESNTDYLYPEQLKMTVVKLISHREC 190
Db 163 SRCQAESVVRTVCLPPADLQPLDPTWTECELSGYGKEALSPFYSERLKEAHRVLPSSRC 222
QY 191 QPPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLOGRMTLTGIVSWG 244
Db 223 TSOHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLISWGLG 282
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
Db 283 CGQKDPGVYTKVTNYLDWIRDNR 307

RESULT 23
US-09-987-455-17
; Sequence 17, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tavapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
```

```
; SEQ ID NO 17
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: K2S 220-527
US-09-987-455-17

Query Match          38.7%; Score 583; DB 11; Length 308;
Best Local Similarity 44.9%; Pred. No. 3e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

Qy 13 CG-QKTLRPRFKIIGGEFTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 45 CGLRQYSQPFRIKIGLGFADIASHPWQAIFAHRSPGGERFLCGGILISSCWILSAHC 104

Qy 71 FIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 162

Qy 131 GRCAQPSRTIOTICLPMSYNDPQFOTSCBITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 163 SRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLYPSSRC 222

Qy 191 QQPHYVGSVTTKMLCAAD----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 223 TSQHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 282

Qy 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 CGQKDPGVYTKVTNYLDWIRDNR 307

RESULT 24
US-09-987-457-11
; Sequence 11, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: part of the
US-09-987-457-11

Query Match          38.7%; Score 583; DB 11; Length 331;
Best Local Similarity 44.9%; Pred. No. 3.3e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

Qy 13 CG-QKTLRPRFKIIGGEFTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 CGLRQYSQPFRIKIGLGFADIASHPWQAIFAHRSPGGERFLCGGILISSCWILSAHC 127

Qy 71 FIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 185

Qy 191 QQPHYVGSVTTKMLCAAD----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 CGQKDPGVYTKVTNYLDWIRDNR 330

RESULT 26
US-09-987-457-13
; Sequence 13, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
```

```
Qy 131 GRCAQPSRTIOTICLPMSYNDPQFOTSCBITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 SRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLYPSSRC 245

Qy 191 QQPHYVGSVTTKMLCAAD----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 TSQHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 305

Qy 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 CGQKDPGVYTKVTNYLDWIRDNR 330

RESULT 25
US-09-987-455-12
; Sequence 12, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: K2S 197-527
US-09-987-455-12

Query Match          38.7%; Score 583; DB 11; Length 331;
Best Local Similarity 44.9%; Pred. No. 3.3e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

Qy 13 CG-QKTLRPRFKIIGGEFTTIENQWFAAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 CGLRQYSQPFRIKIGLGFADIASHPWQAIFAHRSPGGERFLCGGILISSCWILSAHC 127

Qy 71 FIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 185

Qy 131 GRCAQPSRTIOTICLPMSYNDPQFOTSCBITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 SRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLYPSSRC 245

Qy 191 QQPHYVGSVTTKMLCAAD----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 TSQHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 305

Qy 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 CGQKDPGVYTKVTNYLDWIRDNR 330

RESULT 26
US-09-987-457-13
; Sequence 13, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
```



```
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-13

Query Match          38.7%; Score 583; DB 11; Length 335;
Best Local Similarity 44.9%; Pred. No. 3.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-OXTLRPRFKIIGGEFTTIENQWPFAAIYRRH-RGGSVTVVCGGSLISPCWVISATHC 70
DB 72 CGLRQVSPQPRFKIGGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHC 131
QY 71 FIDYPKKEDYIVYGRSLNNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSK 130
DB 132 FQERFPFPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDS 189
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFKENSTDYLYPEQLKMTVVKLISHREC 190
DB 190 SRCQAQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 249
QY 191 QOPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGSGPLVCSLOGRMTLTGIVSWG 244
DB 250 TSQHLLNRTVTDNMLCAGDTRSGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 309
QY 245 CALKQKPGVYTRVSHFLPWIRSHTK 269
DB 310 CGQKQVPGVYTKVTNYLDWIRDNR 334

RESULT 27
US-09-987-455-14
; Sequence 14, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-12

Query Match          38.7%; Score 583; DB 11; Length 339;
Best Local Similarity 44.9%; Pred. No. 3.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-OXTLRPRFKIIGGEFTTIENQWPFAAIYRRH-RGGSVTVVCGGSLISPCWVISATHC 70
DB 76 CGLRQVSPQPRFKIGGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHC 135
QY 71 FIDYPKKEDYIVYGRSLNNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSK 130
DB 136 FQERFPFPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDS 193
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFKENSTDYLYPEQLKMTVVKLISHREC 190
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 193-527,
; OTHER INFORMATION: modified
US-09-987-455-14

Query Match          38.7%; Score 583; DB 11; Length 335;
Best Local Similarity 44.9%; Pred. No. 3.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-OXTLRPRFKIIGGEFTTIENQWPFAAIYRRH-RGGSVTVVCGGSLISPCWVISATHC 70
DB 72 CGLRQVSPQPRFKIGGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHC 131
QY 71 FIDYPKKEDYIVYGRSLNNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSK 130
DB 132 FQERFPFPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDS 189
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFKENSTDYLYPEQLKMTVVKLISHREC 190
DB 190 SRCQAQESSVVRTVCLPPADLQLPDWTCELSGYGKHEALSPFYSERLKEAHVRLYPSRC 249
QY 191 QOPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGSGPLVCSLOGRMTLTGIVSWG 244
DB 250 TSQHLLNRTVTDNMLCAGDTRSGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 309
QY 245 CALKQKPGVYTRVSHFLPWIRSHTK 269
DB 310 CGQKQVPGVYTKVTNYLDWIRDNR 334

RESULT 28
US-09-987-457-12
; Sequence 12, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-12

Query Match          38.7%; Score 583; DB 11; Length 339;
Best Local Similarity 44.9%; Pred. No. 3.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-OXTLRPRFKIIGGEFTTIENQWPFAAIYRRH-RGGSVTVVCGGSLISPCWVISATHC 70
DB 76 CGLRQVSPQPRFKIGGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAAHC 135
QY 71 FIDYPKKEDYIVYGRSLNNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSK 130
DB 136 FQERFPFPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YNDNIALQLKSDS 193
QY 131 GRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFKENSTDYLYPEQLKMTVVKLISHREC 190
```

```

; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tavapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652-2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-14

```

RESULT 29

```

US-09-987-455-13
; Sequence 13, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tavapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 002779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 193-527,
; OTHER INFORMATION: modified
US-09-987-455-13

```

```

Query Match 38.7%; Score 583; DB 11; Length 339;
Best Local Similarity 44.9%; Pred. No. 3.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFAPAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
DB 76 CGLRQYQPFRIKGGFLFADIAHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 135
QY 71 FIDYPKKEDIYVILGRSLNSNTQGMKFVENILHDKYSADTLAHNDIALKIRSK 130
DB 136 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YNDIALQLKSDS 193
QY 131 GRCAQPSRTICTLPSMYNDPQGTSCETITGFGKENSTDYLYPEOLKMTVVKLISHREC 190
DB 194 SRCAQESSVVRTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRC 253
QY 191 QQPHYVGSEVTTKMLCAAD-----POWKT-DSQCGDSGGPLVCSLQGRMTLTGIVSWG 244
DB 254 TSQHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDSDGGPLVCLNDGRMTLVGIISWGLG 313
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 314 CGQKDVPGVYTKVTNYLDWIRDNR 338

```

RESULT 30

```

US-09-987-457-14
; Sequence 14, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya

```

```

; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tavapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652-2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-14

Query Match 38.7%; Score 583; DB 11; Length 343;
Best Local Similarity 44.9%; Pred. No. 3.5e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFAPAIYRRH-RGGSVTYVCGSLISPCWVISATHC 70
DB 80 CGLRQYQPFRIKGGFLFADIAHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 139
QY 71 FIDYPKKEDIYVILGRSLNSNTQGMKFVENILHDKYSADTLAHNDIALKIRSK 130
DB 140 FQERFPPHLLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDDT--YNDIALQLKSDS 197
QY 131 GRCAQPSRTICTLPSMYNDPQGTSCETITGFGKENSTDYLYPEOLKMTVVKLISHREC 190
DB 198 SRCAQESSVVRTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRC 257
QY 191 QQPHYVGSEVTTKMLCAAD-----POWKT-DSQCGDSGGPLVCSLQGRMTLTGIVSWG 244
DB 258 TSQHLNRTVTDNMLCAGDTRSGGPGQANLHDACQDSDGGPLVCLNDGRMTLVGIISWGLG 317
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 318 CGQKDVPGVYTKVTNYLDWIRDNR 342

Search completed: January 8, 2004, 13:00:38
Job time : 34 secs

```